ERGONOMICS

UNIVERSITY OF HAWAII
LIBRARY
JUL 12 '61

HUMAN FACTORS IN WORK, MACHINE CONTROL AND EQUIPMENT DESIGN

A Taylor and Francis International Journal
The Official Publication of the Ergonomics Research Society

Volume 4 **Number 2**

April 1961

ALERE FLAMMAM.
Printed and Published by
TAYLOR & FRANCIS LTD.
RED LION COURT, FLEET STREET, LONDON, E.C.4

ERGONOMICS

Human Factors in Work, Machine Control and Equipment Design

General Editor
A. T. Welford
St. John's College, Cambridge

Associate Editor (Anatomy and Physiology)
W. F. Floyd
Department of Ergonomics and Cybernetics, Loughborough College of Technology, Loughborough, Leicestershire.

Reviews Editor
W. T. Singleton
College of Aeronautics, Cranfield, Bletchley, Buckinghamshire.

Assistant Editor
Miss H. M. Clay
Department of Scientific and Industrial Research, Cornwall Terrace, London, N.W.1.

Editorial Board

Belgium: Dr. H. Bastenier. Laboratoire d'Hygiène, Université Libre de Bruxelles.
Canada: Dr. R. B. Bromiley. Defense Research Medical Laboratories, Post Office Box 62, Postal Station "K", Toronto 12, Ontario.
France: Professor R. Bonnardel. École Pratique des Hautes Études, Laboratoire de Psychologie Appliquée, 41 Rue Gay-Lussac, Paris 5.
Professor Bernard Metz. Institut de Physiologie, Faculté de Médecine, Université de Strasbourg.
Germany: Professor E. A. Müller. Max-Planck-Institut für Arbeitsphysiologie, Dortmund.
Great Britain: M. G. Bennett. Research Department, British Railways, 20 Euston Square, London, N.W.1.
Dr. W. E. Hick. University of Cambridge, Psychological Laboratory, Downing Place, Cambridge.
Emeritus Professor Sir Charles Lovatt Evans, Hedgemoor Cottage, Winterslow, Wilts.
Dr. L. G. Norman. London Transport Executive Medical Department, 280 Marylebone Road, London, N.W.1.
Netherlands: Dr. F. H. Bonjer. Nederlands Instituut voor Praeventieve Geneeskunde, Leiden.
Sweden: Professor S. P. M. Forssman. S. Blasieholmshamnen 4A, Stockholm, 16.
Switzerland: Professor E. Grandjean. Institut für Hygiene und Arbeitsphysiologie der Eidgenössischen Technischen Hochschule, Zürich.
U.S.A: Professor H. S. Belding. University of Pittsburgh, Graduate School of Public Health, Pittsburgh 13, Pennsylvania.
Professor P. M. Fitts. Department of Psychology, University of Michigan, Ann Arbor, Michigan.

Price 25s. 0d. per part
Subscription price per volume £4 15s. 0d. post free, payable in advance

Agents for U.S.A. and Canada
Academic Press Inc., 111 Fifth Avenue, New York 3, N.Y., U.S.A.
$3.50 per part. $13.30 per volume.

ENGINEERING
AND
INDUSTRIAL PSYCHOLOGY

Editor: LEE W. COZAN

Quarterly publishing original investigations and occasional theoretical papers dealing with the adaptation of human tasks and the working environment to the psychological and physiological attributes of human beings ; and the application of psychological principles and research methods to the solution of personnel management problems in industry, government and military establishment

Subscription Rate

DOMESTIC: $7.00

OVERSEAS: $8.00

SINGLE COPIES: $2.00

Orders and Manuscripts should be addressed to:

ENGINEERING AND INDUSTRIAL PSYCHOLOGY
Postoffice Box 662
Washington 4, D. C.

Immediate and Low Level Effects of Ionizing Radiations

Edited by A. A. Buzzati-Traverso

Proceedings of The Symposium held at Venice in June 1959 under the joint sponsorship of U.N.E.S.C.O., I.A.E.A. and C.N.R.N. and published as a supplement to the International Journal of Radiation Biology

The effects of low level radiation, from a biological point of view, have received relatively little attention during many meetings on radiobiology held in the past years. The subject, however, appears of primary significance at this time, in view of the widespread use of radiation sources for scientific, medical, industrial and military purposes. For this reason a symposium exclusively devoted to the subject was called for. Thanks to the financial support of U.N.E.S.C.O., I.A.E.A. and the Italian Atomic Energy Agency (C.N.R.N.) and the hospitality of the Fondazione Giorgio Cini, Z. M. Bacq (Belgium), E. Boeri (Italy), A. A. Buzzati-Traverso (Italy) and A. Hollaender (U.S.A.) organized an international symposium on " The Immediate and Low Level Effects of Ionizing Radiations " which has been held at Venice in June 1959. The meeting was attended by some 120 specialists from many countries.

Price—£2 15*s.* 0*d.* ($8.00) plus postage

Printed and Published by

TAYLOR & FRANCIS LTD
RED LION COURT, FLEET STREET, LONDON, E.C 4

REVISED CHEAPER EDITION FOR LIBRARIES AND SCHOOLS

A History of Mathematics

From antiquity to the early nineteenth century

By J. F. SCOTT, B.A., D.Sc., Ph.D.

Vice-Principal of St. Mary's College, Strawberry Hill, Twickenham, Middlesex

Author of *The Scientific Work of René Descartes* (1596–1650),
Mathematical Work of John Wallis, D.D., F.R.S. (1616–1703), and other works

CONTENTS : Mathematics in Antiquity—Greek Mathematics—The Invention of Trigonometry—Decline of Alexandrian Science and the Revival in Europe—Mathematics in the Orient—Progress of Mathematics during the Renaissance—New Methods in Geometry—The Rise of Mechanics—The Invention of Decimal Fractions and of Logarithms—Newton and the Calculus—Taylor and Maclaurin, the Bernoullis and Euler, Related Advances—The Calculus of Variations, Probability, Projective Geometry, Non-Euclidean Geometry—Theory of Numbers—Lagrange, Legendre, Laplace, Gauss.

This volume is intended primarily to help students who desire to have a knowledge of the development of the subject but who have too little leisure to consult works and documents. The author has availed himself of the facilities afforded by the Royal Society and other learned Societies to reproduce extracts from manuscripts and many scarce works.
Size 9¾″×6¾″. 266 pp. **Price 27s. 6d.,** plus postage and packing 2*s.* 0*d.*

Some Reviews of the First Edition

" The invention of trigonometry, decimal fractions, logarithms and the calculus are each discussed clearly and concisely. The book is easy to read for anybody who knows the elements of mathematics and, although not free from minor errors, can be strongly recommended."—*British Book News*, April 1958.

" Physicists will find that the development in applied mathematics are clearly set out, from ancient times, through that of Archimedes, to the mechanics of the sixteenth century when interest was revived. Significant advances made by Stevin, Galileo, Descartes, Huygens and others are stressed, and help the reader to appreciate what Newton achieved. There are useful appendices giving brief biographical notes on mathematical topics and terminology, followed by a bibliography."—*Proceedings of The Physical Society*, September 1958.

" . . . it has been written with clarity and balance, and the excellent printing helps to make it a pleasure to read."—*The Times Educational Supplement*, 21 March 1958.

". . . his (Dr. Scott's) wide knowledge of the material, his careful description of methods combine to provide an account which at times gives a sense of excitement of discovery,"—*Nature*, 26th July 1958.

" The printers and publishers are to be congratulated upon having produced such an attractive volume, ... We feel sure the book will be received with delight by all those interested in mathematical histories."—*BEAMA Journal*, August 1958.

" The work comes to life mainly because of his admirable use of the writings of mathematicians themselves, which vividly illustrates the great difficulties under which many of them laboured. This is not a book for the layman but both the student and anyone to whom figures are a fascination, will find the subject clearly and pleasantly presented."—*Technical Bookguide*, March 1958.

Printed and Published by

TAYLOR & FRANCIS LTD
RED LION COURT, FLEET STREET, LONDON, E.C.4

A PRACTICAL METHOD OF ESTIMATING AN INDIVIDUAL'S MAXIMAL OXYGEN INTAKE

By J. S. Maritz, J. F. Morrison, J. Peter, N. B. Strydom and C. H. Wyndham

Applied Physiology Laboratory,
Transvaal Chamber of Mines Research Laboratories, Johannesburg

The premises upon which simple methods for estimating an individual's maximum oxygen intake are based have been tested. These premises are :

(*a*) that heart rate and oxygen intake are linear functions of each other throughout the entire range of work up to the individual's maximum ;

(*b*) that oxygen intake of the individual deviates very little from the mean straight line relating oxygen intake and rate of work for the population, so that the oxygen intake for a task performed against gravity can be estimated with reasonable precision from the rate of work ; and

(*c*) that the individual variability of maximum heart rate round the mean for the population is sufficiently small to use the mean in a routine test procedure without the introduction of large errors.

Premises (*b*) and (*c*) are fully substantiated by the results reported here. Premise (*a*) is not strictly valid, as there is a bias of the order of 0·3 l. oxygen/min in the straight line relating heart rate and oxygen intake at high levels of work in most individuals.

The errors involved in the method proposed by Åstrand (1954) based on these premises have been examined and compared with those of an alternative procedure described in the present paper. The variance of maximum oxygen intake using the Åstrand method, due to errors in measurement of heart rate alone, is of the order 0·53, compared with 0·26 of the procedure described in this paper. In addition, one measurement only of heart rate in the Åstrand method introduces a bias of about 0·3 l. oxygen/min in the estimate of an individual maximum oxygen capacity of 3·0 l./min. The variance of maximum oxygen intake by the present method can be reduced by fitting a straight line by least squares to plots of four pairs of heart rate and oxygen intake values at four rates of work.

§ 1. Introduction

Most manual labour is carried out on aerobic metabolism—that is to say, the rate of oxygen utilization for the chemical processes concerned with the release of energy in muscle is balanced by the rate of oxygen intake. The physical effort required for the period of an average shift could not be sustained otherwise. Quantitative expression has been given to these considerations by experienced research workers in this field such as Christensen (1953), Brody (1945), Åstrand (1952) and others. Their opinion is that the average level of effort for endurance work should not require more than half the maximum oxygen intake of the individual.

An individual's maximum oxygen intake can be determined by the method proposed by Åstrand (1952) or by that of Taylor *et al.* (1955). These two test procedures differ in detail but are otherwise similar. The subject exercises at successively higher rates of work until a level is reached at which : further increase in the rate of work does not produce a significant increase in rate of oxygen intake ; the time for the onset of acute fatigue is short ; and the concentration of lactic acid in the blood rises sharply. These methods are tedious and time-consuming and, moreover, the present criteria proposed by these workers for assessing accurately the 'true' level of maximum oxygen intake by laboratory methods are not acceptable, as pointed out in studies by Wyndham *et al.* (1959).

The S. African gold mining industry recruits annually a labour force of about 300 000 untrained natives of Southern Africa and, because it is migrant, labour turnover is almost 100 per cent per annum. A relatively large proportion of the mining operations are carried out with manual labour. Under these circumstances the measurement of the maximum oxygen intake of an individual recruit has potential value in placing the right man with the minimum of delay in the right place as far as his capacity for physical effort is concerned. The laboratory procedures of Taylor *et al.* (1955) and Åstrand (1952) are too lengthy and costly in terms of scientific staff and apparatus to apply to a labour force of this size with a high rate of annual turnover.

An alternative to the laboratory test, which has the merit of being simple and brief, is the procedure proposed by Åstrand and Rhyming (1954). The maximum level of oxygen intake of an individual is estimated from a nomogram, using the measurement of heart rate and prediction of the rate of oxygen intake from the work rate at a SINGLE level of step climbing. The nomogram is based upon certain premises, implied but not explicitly stated. The first premise is that heart rate is a linear function of oxygen intake up to maximum rates of work ; the second, that the deviations of individual maximum heart rates from the mean maximum for the population are small ; and the third is that all the individual straight lines representing the relationship between heart rate and oxygen intake have a common point which for males is 60 beats/min at zero levels of oxygen intake. A fourth premise is that the oxygen intake of the individual deviates very little from the mean straight line relating oxygen intake to rate of work for the population, so that the oxygen intake for a task performed against gravity, such as climbing on and off a step, can be estimated with reasonable precision from the rate of work. For step climbing the work rate can be calculated from the product of gross body weight, the height of the step and the rate of climbing. (This generalization fails to take account of energy losses over the whole cycle of stepping, but for the present purpose is exact enough.)

If these premises are valid, then the following procedure, expressed in nomogram form in the paper by Åstrand and Rhyming (1954), can be used to assess the individual maximum level of oxygen intake with reasonable precision :—

> Heart rate is measured at a certain rate of step climbing and from the average oxygen intake/work rate graph the corresponding oxygen intake is estimated. This is then plotted against the observed heart rate, giving a single point on an oxygen intake/heart rate grid. Through this point and the ‘ common point ’ a straight line is drawn and extrapolated to a heart rate of 128 or 196 beats/min, and the corresponding oxygen intake is an estimate of the half maximum or of the maximum oxygen intake capacity.

From the point of view of simplicity and time Åstrand and Rhyming's procedure has much to commend it. However, in our view the test cannot be applied immediately in industry because some of the techniques are not acceptable and certain premises do not appear to have been tested rigidly enough. The unacceptable techniques are the fitting of the straight line through the heart rate of 60 beats/min at the zero oxygen intake and the plot of a SINGLE measurement of heart rate at ONE rate of work only. Heart rate determinations are subject to ' individual ' random variation due to ' error of measurement ' and ' natural physiological ' causes. Hence a straight line fitted through two points, one of which is an assumed and perhaps biased value and the other subject to these variations, may have a considerable error. The error is reduced if three, or preferably four, measurements are made and especially if a straight line is fitted by the method of least squares.

In order to test the premises on which the test procedure rests, we have made a number of studies to establish :—

(*a*) The form of the functional relationship between heart rate (H) and work rate (W), between oxygen intake (O) and W, and between H and O, under the following headings :—

- (i) whether the heart rate and oxygen intake are, to a close degree of approximation except for random variation, linear functions of the rate of work ;
- (ii) whether heart rate and oxygen intake reach some asymptotic maximum value at high levels of work ;
- (iii) whether, if (i) and (ii) are correct, heart rate and oxygen intake are linear functions of each other throughout the entire range of work rate up to an individual's maximum.

(*b*) The feasibility, in a routine scheme for testing recruits, of substituting for the actual measurements of oxygen intake values read off a mean population graph relating oxygen intake to work rate.

(*c*) Whether the individual variability of maximum heart rate values round the MEAN for the population is sufficiently small to use the population mean in a routine test procedure without the introduction of large errors.

It is important to draw attention to the fact that these studies were made in Johannesburg, at an altitude of 6000 ft above sea level.

§ 2. METHODS

2.1. *Design*

Four African mine labourers recruited from one of the Shangaan tribes in Mozambique were trained to step on and off a stool 1 ft in height at a rate of 24 steps/min. When they were trained a series of studies were made which lasted five months. On different days of the week, from Monday to Friday, they stepped at 6 steps/min, 12 steps/min and 24 steps/min for time periods which varied from two to six hours. The rate of work on each Monday was 24 steps/min. For the remainder of the week the above rates were taken at random. A few tests were made at a rate of 30 steps/min.

At the end of the above study the four subjects were trained to pedal at 70 rev/min on a bicycle ergometer. When trained to this procedure, they worked for a total period of four months at work loads equivalent to 2500, 5000, 7000, 8000, 8500, 9000, 9500, 10 000, 10 500 and 11 000 ft lbs/min. On each Monday morning the rate of work was 7000 ft lbs/min. On other days of the week the rate of work was taken at random.

A group of 26 African mine recruits, chosen at random, was studied on recruitment to one of the mines and again one month later and a further four months after recruitment. The initial physiological measurements were made before breakfast while the men rested in bed, having slept the night in a hospital ward. During the morning they performed three half-hour periods of step climbing. The work periods followed each other without pause and, in succession, were 6, 12 and 24 steps/min.

2.2. *Observations*

Heart rate was determined during the initial few months by means of a stethoscope at the apex of the heart and was counted over half a minute. Subsequently, measurements were made with a portable Sanborn electrocardiograph.

Oxygen consumption was measured by the conventional Douglas bag and Haldane gas analysis method. Duplicate aliquot samples of expired air taken for analysis were required to agree to within 0·05 vol. per cent. At low levels of work gas collections generally were over three min, but at the very high levels on the bicycle ergometer collections had to be limited to 1 min.

2.3. *Subjects*

Physical characteristics of the four trained African recruits from Mozambique are given in Table 1. The height, weight and tribe of the 26 recruits chosen at random are given in Table 2. Accurate ages of these men cannot be given as their births are not registered. They are all young male adults in the early 20's.

Table 1. Physical characteristics of four trained men

	On arrival		End of series	
No.	Height (in.)	Weight (lb oz)	Height (in.)	Weight (lb oz)
1	$62\frac{1}{4}$	144·5	$62\frac{1}{2}$	146·2
2	$61\frac{3}{4}$	121·10	$61\frac{3}{4}$	118·11
3	64	125·1	64	125·8
4	$64\frac{3}{4}$	127·4	$64\frac{3}{4}$	131·8

2.4. *Apparatus*

At high levels of work the ventilation rate of the subjects reached 150 litres/min, and respiratory valves which would offer little resistance to rates of flow of this order were therefore used.

The bicycle ergometer, electromagnetically braked, was calibrated at approximately monthly intervals. To ensure consistency of work rate the pedalling rate was recorded electrically with a generator and galvanometer-type recorder ; the times of gas collection were marked on the record.

§ 3. RESULTS

3.1. *Examination of the relationship between H and W, between O and W, and between H and O*

3.1.1. *Repeated observations on four trained men at submaximal work loads*

A large number of measurements of H and O were made over a period of months on four trained subjects while they were stepping on and off the 1 ft stool at rates of 6, 12, 24 and 30 steps/min. Direct estimates of residua variances of H and O could therefore be made and used to test whether a straight line is a good fit to plots of H/W, O/W and H/O, i.e. in order to test the first premise (a) outlined in Section 1.

Table 2. Physical characteristics of 26 recruits

No.	Height (in.)	Weight (lb oz)	Tribe
7	66	125·14	—
8	68	116·2	Basuto
9	66½	116·5	Shangaan
10	65	118·2	Shangaan
11	68½	112·11	Shangaan
12	63	115·6	Basuto
13	70	142·13	Hlubi
14	68	110·11	Bechuana
15	64	110·10	Xosa
16	72½	128·12	—
17	70½	135·11	Bechuana
18	61½	115·5	Xosa
19	62½	120·0	Xosa
20	65½	131·2	Basuto
21	64½	107·12	Xosa
22	65¾	110·8	Basuto
23	67	127·14	Xosa
24	67	114·12	Xosa
25	63	112·5	Bechuana
26	70¼	119·5	Bechuana
27	67	150·7	Basuto
28	65	128·5	Basuto
29	64	117·12	Hlubi
30	67	123·12	Basuto
31	66	117·12	Shangaan
32	67½	122·0	Xosa

Table 3. Observed mean values of H and O and variances

	No. 1			No. 2			No. 3			No. 4		
Work rate	N	Mean	Res. Var.	N	Mean	Res. Var.	N	Mean	Res. Var.	N	Mean	Res. Var.
'H'												
6×12	5	73·6	84·8	5	84·8	123·2	4	70·0	5·3	4	88·0	42·7
12×12	8	88·0	64·0	8	94·5	109·9	3	78·7	37·3	3	104·0	48·0
24×12	31	110·8	53·5	39	113·7	97·6	22	96·6	22·6	22	123·6	42·5
30×12	6	115·3	34·6	6	127·3	73·0	3	98·7	101·3	3	118·7	69·3
'O'												
6×12	3	0·560	0·0018		—	—	3	0·580	0·0108	2	0·559	0·0009
12×12	4	0·943	0·0045	5	0·832	0·0069	2	0·670	0·0102	3	0·767	0·0106
24×12	11	1·455	0·0168	13	1·275	0·0198	4	1·076	0·0275	6	1·181	0·0039
30×12		1·419	0·0258		—	—		—	—	2	1·164	0·0017

The observed means and variances of H and O at the different rates of W for each individual are given in Table 3. The results at 30 steps/min were rejected because the mean values of H and O are no greater than those at 24 steps/min and this indicates that the men were not able to maintain this high value of W.

Straight lines were fitted through the observations of H against W and O against W by the method of least squares. It gives estimates $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ of the parameters α and β in the linear relationships

$x = \alpha + \beta W$: where x is the unknown H or O and
W has the designation as in the text.

The values of α and β for each subject and their standard errors (s.e.) are given in Table 4. The straight lines were tested for goodness of fit by means of the variance-ratio test, resulting in the Snedecor ' F ' as test criterion. Values of F, and the probability of their being exceeded if the hypothesis (of linearity) under test is true, are given in Table 5 and they indicate that, on the whole, the straight lines fit satisfactorily.

Table 4. Estimates of parameters of straight lines fitted to H, W and O, W

	' H, W '				' O, W '			
No.	a_1	s.e.(a_1)	β_1	s.e.(β_1)	a^3	s.e.(a^3)	β^3	s.e.(β^3)
1	65·298	3·174	0·07652	0·006056	0·6420	0·801	0·001273	0·002963
2	75·109	3·912	0·08076	0·008745	0·5117	0·715	0·001552	0·003567
3	63·848	2·952	0·06340	0·006462	0·3892	0·624	0·001362	0·003301
4	85·961	4·511	0·07086	0·009717	0·4229	0·650	0·001378	0·003244

Table 5. Test of goodness of fit based on Snedecor ' F ' value, together with probability (P) of given value being exceeded

	' H, W '			' O, W '		
No.	F	d.f.	P	F	d.f.	P
1	2·434	2, 46	$>0·10$	2·436	4, 17	$>0·10$
2	0·400	2, 54	$>0·50$	4·416	3, 13	$>0·05$
3	1·976	2, 28	$>0·10$	0·439	1, 16	$>0·50$
4	7·854	2, 28	$<0·01$	0·000	1, 8	$>0·50$

As the O, W and H, W relationships are linear the same must be true of the O, H relationship, and its constants can be computed from those of the O, W and H, W straight lines. These straight lines with their 90 per cent confidence bands are given in Fig. 1. The limits are very wide at the high levels of H because there were few observed H values above 140 beats/min.

3.1.2. *One measurement of H and O on* 26 *men at rest and at three rates of W*

H was measured at 5 min intervals and O once during rest and during a 30 min spell of work at each of three values of W, viz., 6, 12 and 24 steps/min. Through each set of four points (plots of H or O against W) a straight line was fitted by least squares. By taking differences between observed and fitted values, squaring, adding and dividing by the appropriate number of degrees of freedom (two in this case), an estimate of the residual variance, based on two degrees of freedom, was found for each of the 26 men. The means of these residual variances are given in Table 6 for comparison with the residual variances found on the four trained men.

Table 6. Mean residual variances of 26 men

$H = 23·1$
$O = 0·00784$

Tests of goodness of fit similar to those for the four men could not be carried out but, if H and O are both linearly related to W, they are also linearly related to each other. Therefore inspection of plots of H against O should enable one to make a reasonable judgment as to the validity of the hypothesis

Figure 1. Straight line, with confidence limits, fitted to plots of heart rates against oxygen intakes for trained subjects Nos. 1–4.

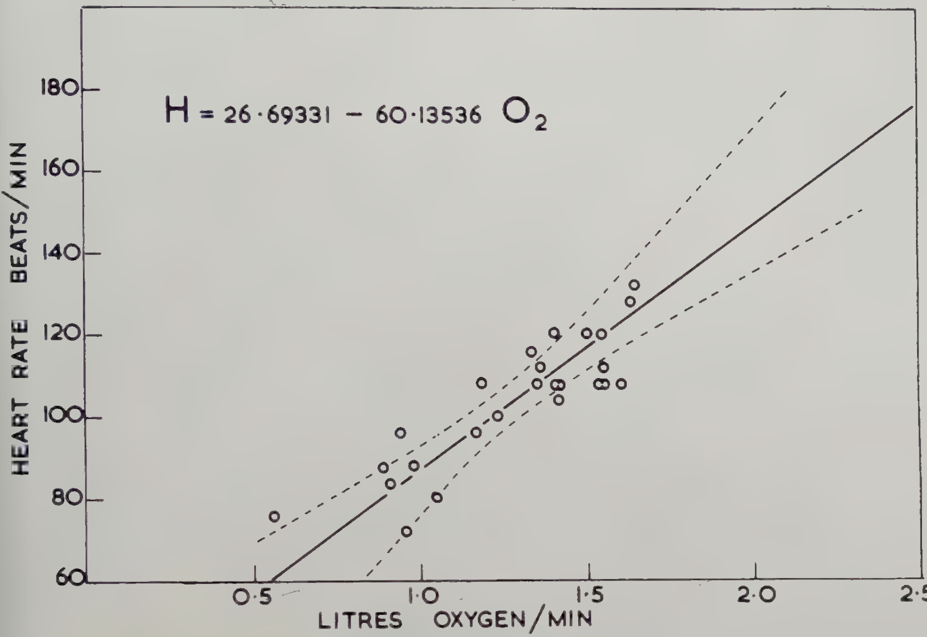


Figure 1. Subject No. 1.

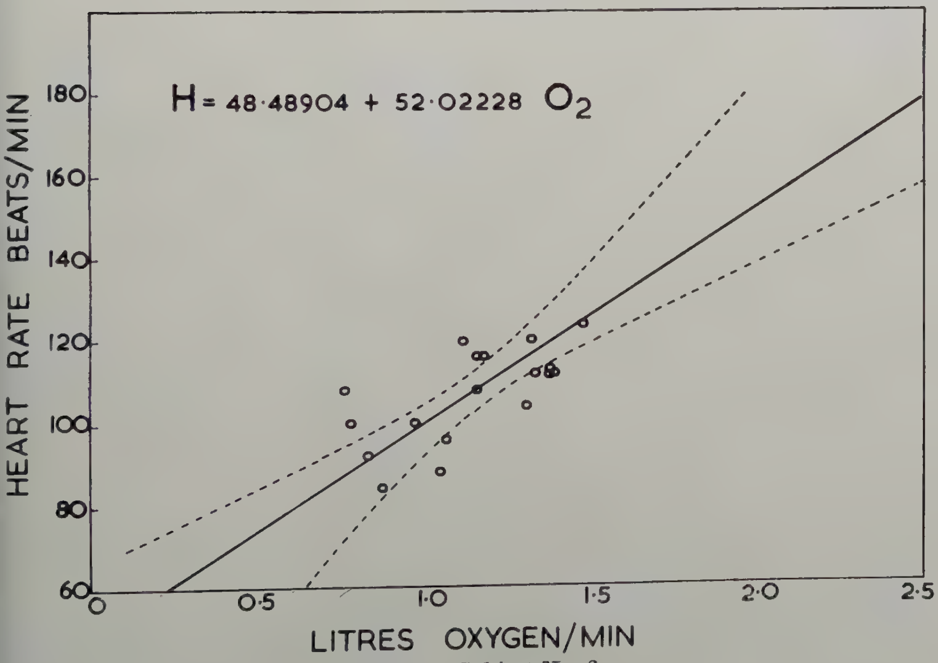


Figure 1. Subject No. 2.

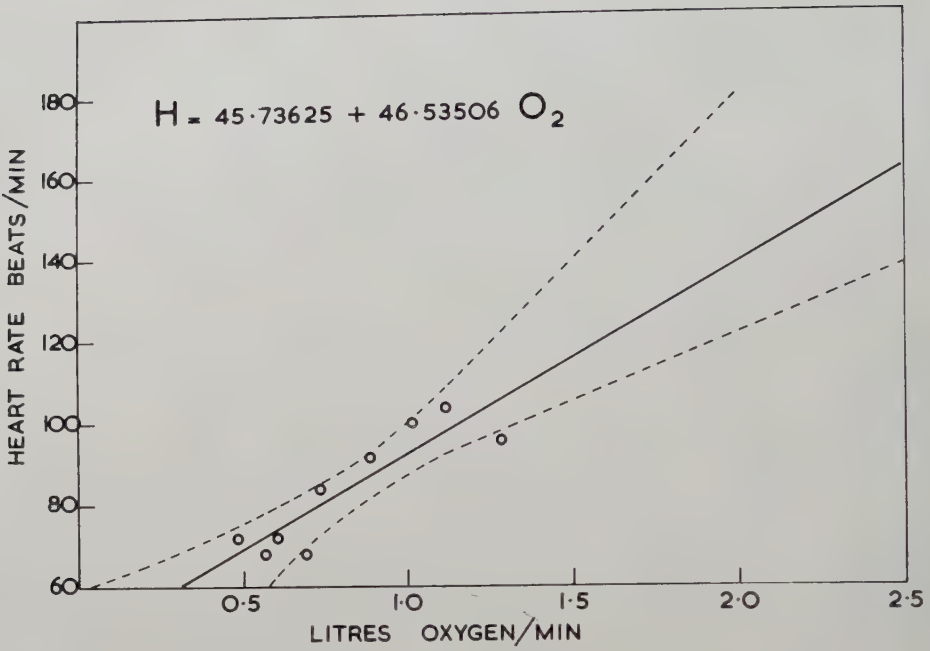


Figure 1. Subject No. 3.

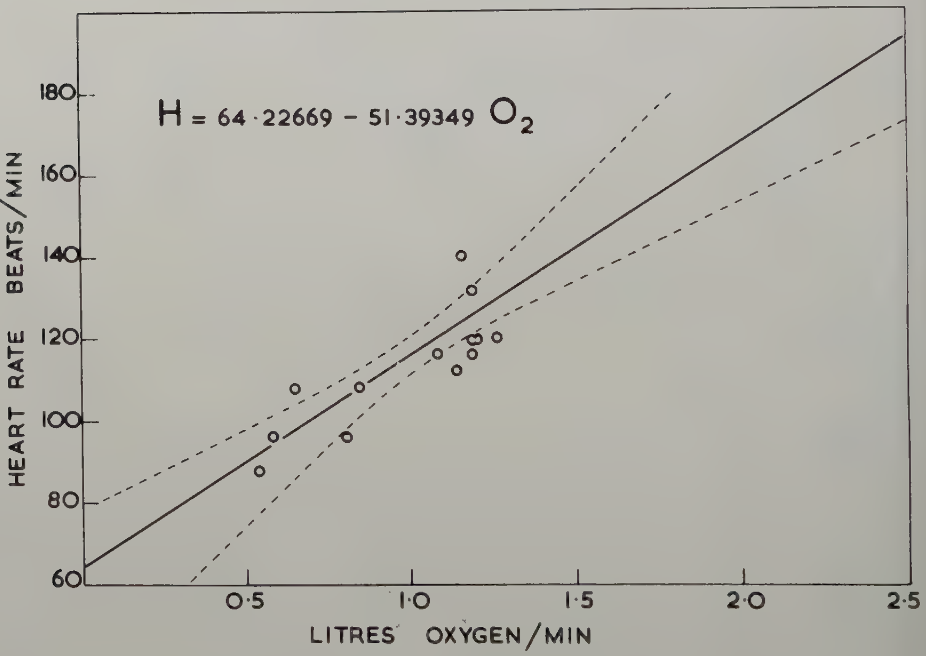


Figure 1. Subject No. 4.

of linearity. Such plots, for each of the 26 men, are given in Fig. 2, and they suggest strongly that the hypothesis is acceptable.

Figure 2 also shows that the straight lines do not have a common point at resting oxygen intakes. An objective test of this observation is available because the variance between subjects of resting values of O and H may be compared with the residual variances given in Table 6. We have the following results :—

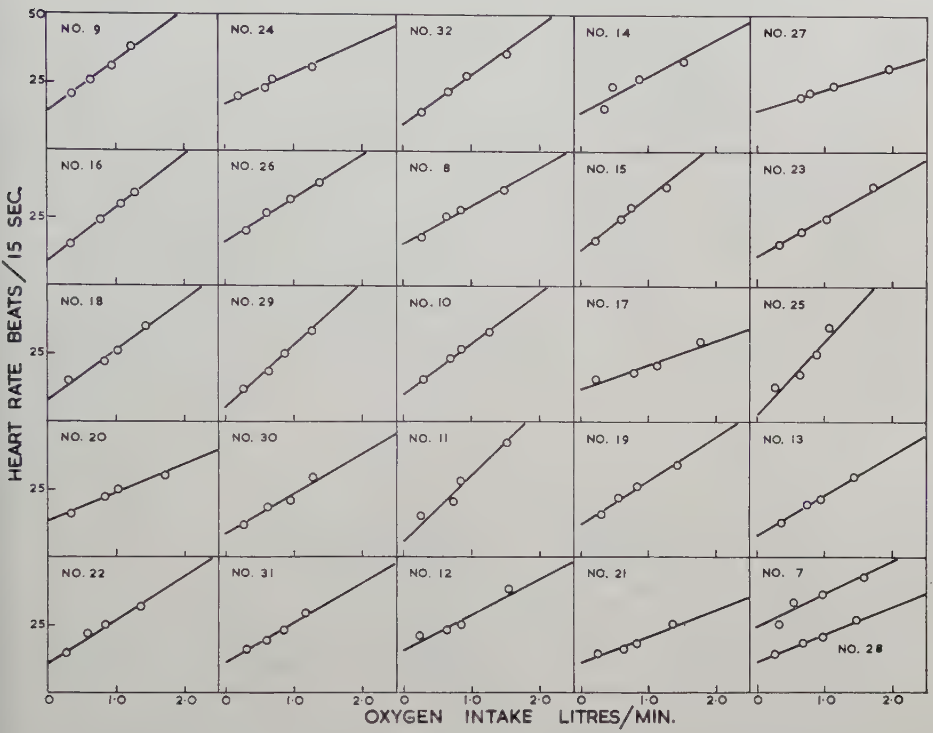


Figure 2. Straight lines fitted to plots of heart rates against oxygen intakes of 26 untrained recruits.

Between subjects variance at resting O $=0{\cdot}00706$.

Between subjects variance at resting $H=160{\cdot}00$ and comparing these with the residual variances we have

$$O : F\begin{pmatrix}25\\25\end{pmatrix}=0{\cdot}90$$

$$H : F\begin{pmatrix}25\\25\end{pmatrix}=6{\cdot}64$$

The differences between resting oxygen intakes are, therefore, not significant, but for resting heart rates they are very significant. This definitely refutes the suggestion that the H, O graphs all have a common point.

3.1.3. *Repeated measurements of H and O at different levels of W up to the individual maximum*

Repeated measurements of H and O were made at a number of values of W up to the individual maximum on the same four trained subjects plus an additional two, using a bicycle ergometer. Direct estimates of residual variances could therefore be made and these could, in turn, be used to test the goodness of fit of the mathematical model. (Ref. Table 11 p. 114.)

The mathematical model used to represent the plots of the data of H against W and O against W, and also the tests for goodness of fit of the model have been detailed in a separate paper by Wyndham *et al.* (1959). It suffices here to state that the model which was found to fit the data well was one with a linear component up to 7750 ft lbs/min level of W and an exponential component above the value of W. The curve relating H to O for each individual was produced by taking a number of values of W up to the individual's maximum and reading off the corresponding H and O values from the respective fitted curves. These ' estimated mean ' values of H and O were then plotted against each other and are given in Fig. 3 for the six trained subjects. At high levels of W there is a small departure from linearity due in most cases to the fact that, as W approaches the individual maximum, O tends to its asymptote more slowly than H. This means that the value of O obtained by extrapolating the straight line to intersect the maximum H value is usually less than the maximum measured O or the value of O at the asymptote. Some idea of the magnitude of the ' bias ' may be gained from Fig. 3, where the straight line

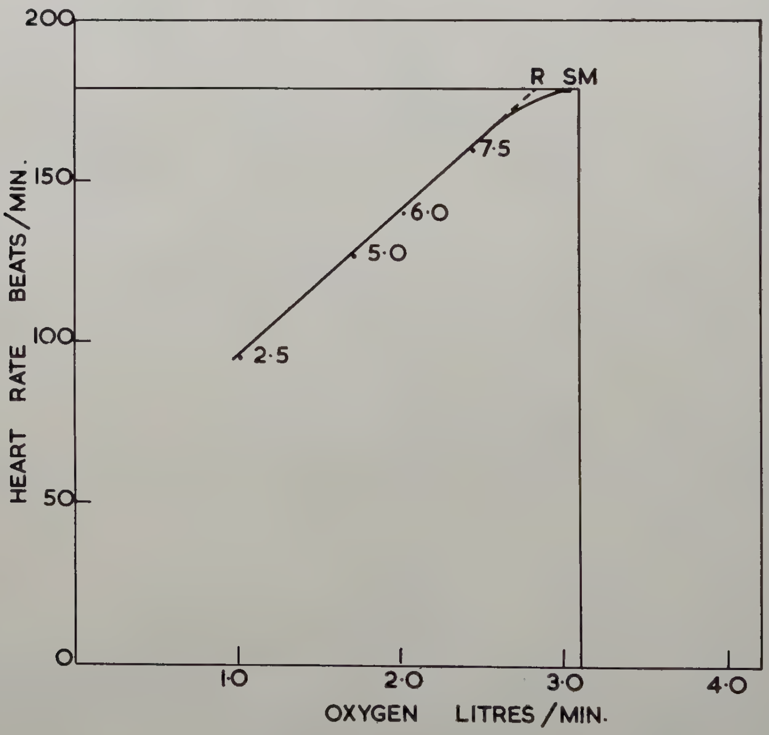


Figure 3. Subject No. 1.

Figure 3. Pairs of heart rates and oxygen intake values between 2500 and 12 000 ft lb/min are taken from the curves fitted to O, W and H, W observations. The above curves are fitted to these pairs of ' mean ' H and O values for subjects Nos. 1–6. Point R is the extrapolation of the straight line part of the fitted curve to the individual's maximum heart rate ; M is the asymptote value of O for the individual's O, W curve ; and S is the mean of the three highest measured values of O. The distances R–M and R–S represent the bias between asymptote (M) and measured values of maximum oxygen intake (S) and that obtained by extrapolating H values at submaximal work loads plotted against O values.

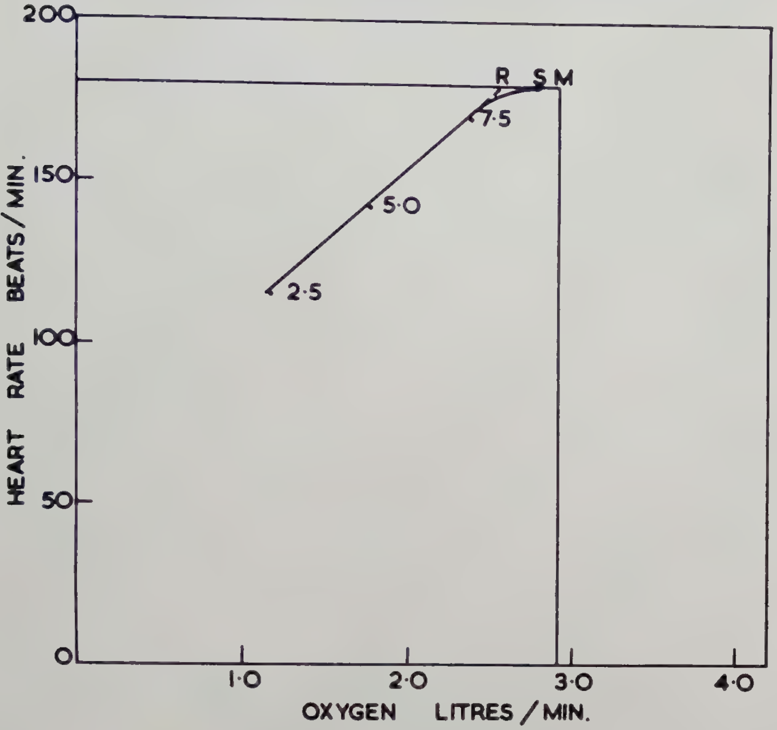


Figure 3. Subject No. 2.

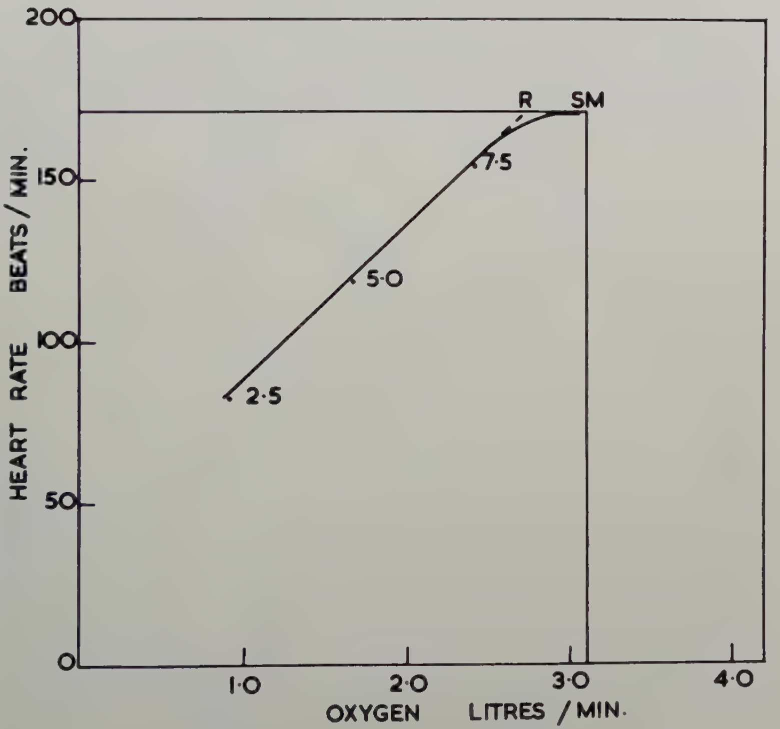


Figure 3. Subject No. 3.

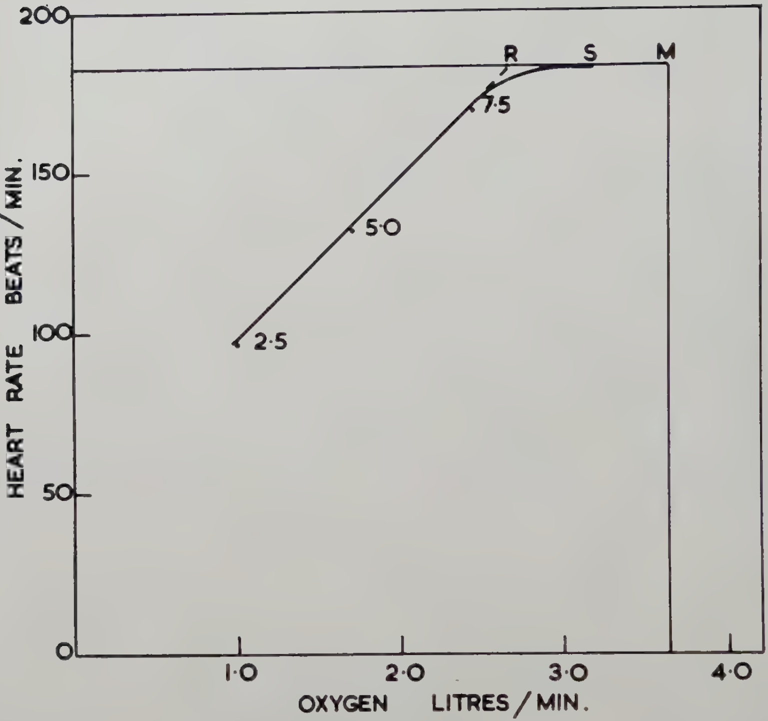


Figure 3. Subject No. 4.

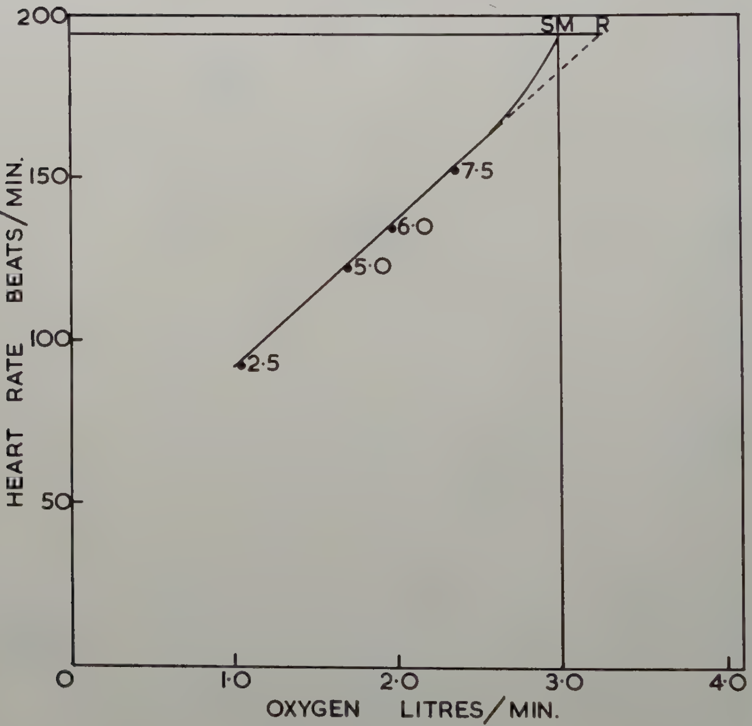


Figure 3. Subject No. 5.

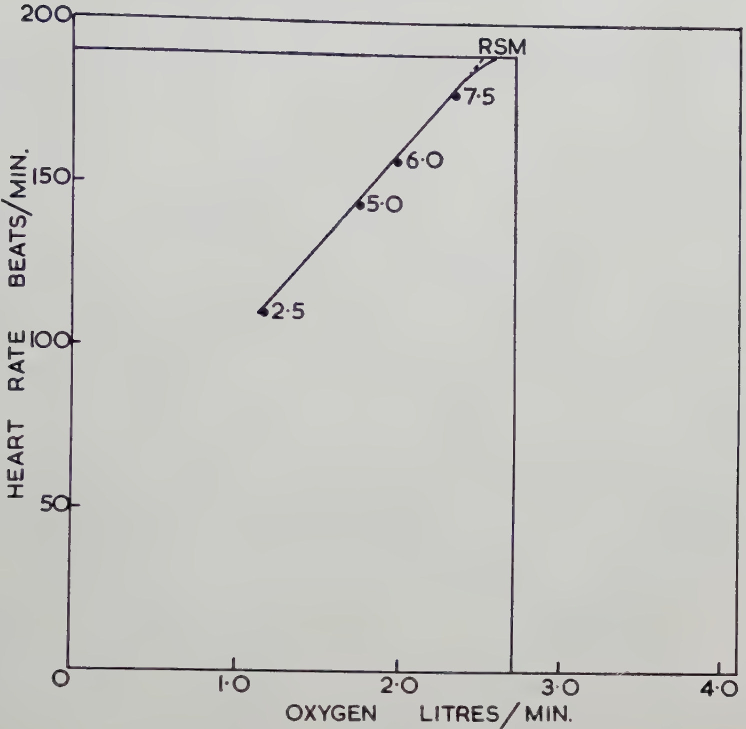


Figure 3. Subject No. 6.

as been extended to intersect the maximum value of H at point R. S correponds to measured O at the highest levels of W and M to the value of O at he asymptote. The distances R–S and R–M are given in Table 7. Two

Table 7. Bias of maximum oxygen intake

No.	$R-S$	$R-M$
1	0·22	0·28
2	0·21	0·35
3	0·33	0·38
4	0·52	0·98
4 (Repeat)	0·101	0·156
5	0·086	0·201
6	−0·260	−1·258

alues are given for subject No. 4. On the first occasion he was tested he did ot exert himself to a level of W at which the asymptote of O corresponded the highest observed values of O. On the second series of experiments one ear later he worked harder and the observed and asymptotic values of O orrespond closely. Both values are given as an illustration of one of the fficulties that can arise in estimating the asymptote of O where a subject es not exert himself to the utmost.

This analysis shows that the relationship between H and O can be reprented by a straight line up to very near the maximum values of W, but that ere is a small ' bias ' towards observed and asymptotic values of O higher an would be obtained by extrapolating the straight line part of the graph the maximum value of H and reading off the corresponding value of O.

3.2. *Mean maximum values and variances of H for a population*

Repeated measurements of H were made at various levels of W up to the individual maximum effort on the six trained men. Appropriate curves were fitted and the goodness of fit tested using the residual variances. Maximum values of H at the asymptote correspond closely to the observed mean maximum values of H at the three highest values of W when at least five observations were made on each individual (Subjects 1, 2, 3, 4, Table 8). When fewer measurements were made (as in the case of Subjects 4 (repeat), 5 and 6, Table 8) the standard error of the means is greater. In the latter the differences between the asymptotic and observed mean maximum H values were greater, although these differences are not significant. The mean asymptotic value of H for the six trained men is 184·1 beats/min with a standard deviation of 8·4.

A further study was made on eight men trained at submaximal work loads on stepping stools but not to maximal loads on the bicycle ergometer. They were required to work up to a rate at which either H on two successive values of W was clearly assuming an exponential form, or at a rate at which they were exhausted within 5 min. These values are also given in Table 8.

Table 8. Comparison of asymptote with measured maximum H

Trained Men	Asymptote H	s.e. of Asymptote	Mean H at 3 highest W's	s.e. of mean H	Untrained men	Mean H at 3 highest W's
1	178·94	1·39	179·29	1·23	No. 33	187
2	180·17	1·87	179·28	1·53	34	177
3	171·09	0·93	171·00	1·54	35	178
4	182·98	1·02	182·26	0·80	36	170
(Repeat	190·63	19·00	174·86	2·67	37	178
5	190·40	9·30	188·12	2·22	38	186
6	194·28	7·26	177·93	1·03	39	184
					40	175
Mean	184·07					179·38

The values of the means of the highly trained and the relatively untrained men are not significantly different, and are so close to 180 beats/min that it may be regarded as a satisfactory working value for young healthy Bantu males.

3.3. *Estimate of rates of oxygen intake*

3.3.1. *Estimate of oxygen intake from rate of work*

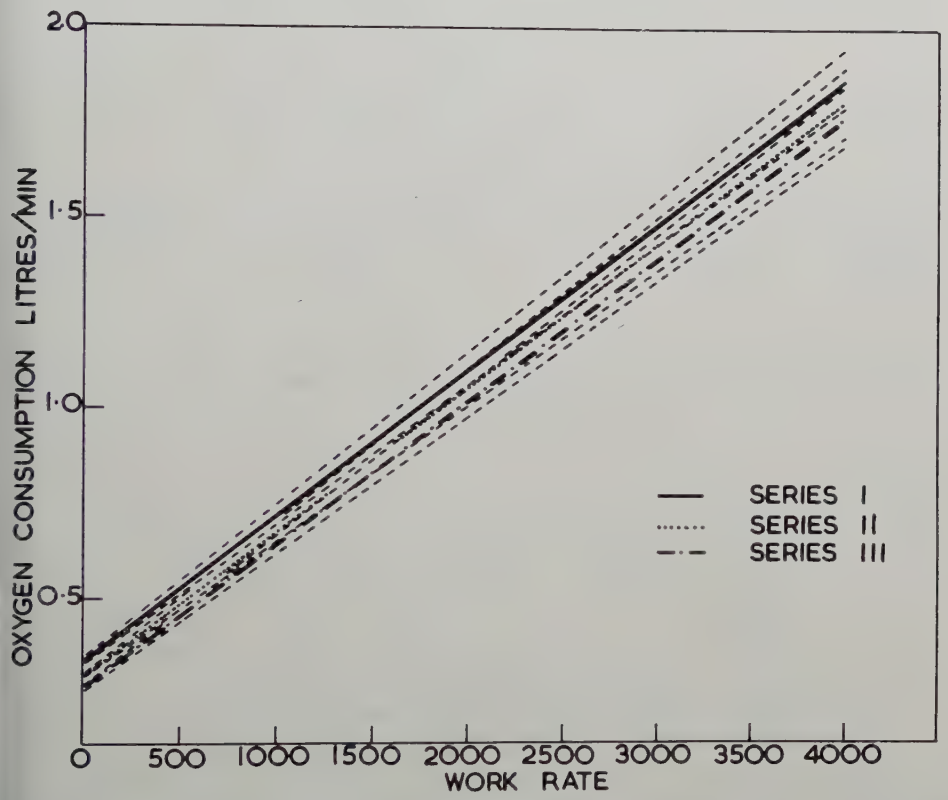
The aim of this part of the study was to examine the magnitude of the individual variations around the population mean O, W line. Thus one could decide whether it would be feasible, in routine practice, to estimate a man's O simply by entering his W in the equation

$$O = A_2 + B_2 W$$

Each individual in the populltion has his own O, W straight line, $O = \alpha_2 + \beta_2 W$, say, and A_2 and B_2 are the population means of α_2's and β_2's respectively. The success of this proposal depends upon the variations between α_2's and β_2's from one individual to another. The random error in estimating O from a 'standard' O, W graph has been calculated (Appendix 6.2) and is less than the variances of measured O due to error and physiological variation, being of the order 0·002 compared with 0·02, the mean of the pooled variances of 'O' of the four subjects (Table 9).

Table 9. Pooled variances of H and O for each subject

	'H'				'O'			
No.	X^2_3	P	Pooled R.V.	d.f.	X^2_3	P	Pooled R.V.	d.f.
1	0·89	$>0·8$	55·5	46	2 20(2)	>0 40	0·0168	17
2	0·33	$>0·95$	98·2	54	1·36(1)	>0 30	0·0166	16
3	5·74	$>0·10$	27·4	28			0·0190	6
4	0·25	$>0·95$	44·8	28			0·0048	9



igure 4. Values for the parameters and in the equation for a straight line, $O = + W$, were obtained from the straight lines fitted to four pairs of O and W values at four submaximal rates of work on 26 recruits. The straight lines, with 95 per cent confidence limits, in this Figure are drawn from the mean of the 26 and values. Series I is on recruitment ; Series II after one month of work ; Series III after a further three months of work.

The above estimates were made on new recruits to the mining industry. efore the procedure could be recommended on a wide scale it was important o see whether the same linear relationship applies in other circumstances, for xample, to men who have been some time at work in the mines. In Fig. 4 given the mean straight line fitted to the individual O, W graphs of the 6 men, together with 90 per cent confidence limits ; in addition there are drawn similar straight lines on O, W data after one month and after a further three months of service. The parameters of the straight lines are so similar that one straight line could be used with little error to estimate O from W for all periods of service.

3.3.2. *Prediction of O from a measurement of ventilation (V)*

Estimation of O from determination of W alone has been considered. If observation of V were used in conjunction with W it might be expected that a more precise estimate of O would be obtained. If incorporation of V resulted in greatly increased precision it would be a worthwhile practical proposition, as its determination does not involve very complicated laboratory procedures.

Calculation of variances of $\boldsymbol{\alpha}_3$ and $\boldsymbol{\beta}_3$ in the eqn.

$$V = \alpha_3 + \beta_3 W$$

enables us to compare standard deviations of O estimated either indirectly from the relationship between V and W or directly from O against W. The standard deviation of the former is approximately 0·9 that of the latter (see Appendix 6.3).

It is clear from this that the improvement in accuracy in measuring V compared with the prediction of O from W would not warrant the extra effort involved.

3.3.3. *Comparison of asymptotic value of O and average O at highest levels of W with O corresponding to maximum H value from extrapolation of H/O line*

From the point of view of the test procedure proposed it is important to compare :

(i) the asymptotic value of O from the O against W curve using the bicycle ergometer data ;
(ii) the mean O at the three highest levels of W from men on the bicycle ergometer ;
(iii) the value of O which corresponds to the maximum H when the straight line fitted to the plots of H against O is extrapolated both for the data obtained on stepping stools and on the bicycle ergometer.

These results are given in Table 10. They show that the O values corresponding to the maximum level of H on the straight line graph of O, W are closely similar whether the bicycle or a step test is employed. The O values are, however, generally less than either the highest measured value of O or the asymptotic value in the O, W curve.

Another feature of the results is the very wide confidence interval of O at the maximum level of H, as determined from the stepping test data. This is due to the very considerable extrapolation of the confidence limits above the highest measured value of H in the tests (about 140 beats/min). This situation could clearly be improved by increasing the work load either by raising the height of the step or adding a weight load to the individual or by increasing the rate of stepping. The last of these alternatives proved impractical, as is seen in Table 3, where values of H and O for 24 and 30 steps/min were not distinguishable.

Table 10. Comparison of maximum oxygen capacities determined by four different procedures

No.	Mean *O* at 3 highest *W*'s (*a*)	*O* at the Asymptote (*b*)	Maximum *O* Bicycle (*c*)	Maximum *O* Step test (*d*)
1	3·01 (2·75–3·27)	3·11 (2·99–3·23)	2·87 (2·60–3·14)	2·55 (2·11–2·99)
2	2·72 (2·62–2·82)	2·90 (2·64–3·16)	2·56 (2·31–2·81)	2·53 (1·97–3·09)
3	3·01 (2·84–3·18)	3·11 (2·97–3·25)	2·92 (2·66–3·18)	2·89 (2·01–3·77)
4	3·13 (2·97–3·29)	3·64 (3·06–4·22)	2·66 (2·57–2·78)	2·25 (1·81–2·69)
Repeat	2·81 (2·59–3·02)	3·06 (2·65–3·46)	2·66 (2·47–2·85)	—
5	2·84 (2·65–3·02)	3·00 (2·74–3·27)	2·93 (2·84–3·02)	—
6	2·55 (2·38–2·72)	2·71 (2·66–2·77)	2·35 (2·25–2·44)	—

§ 4. DISCUSSION

4.1. *Examination of the method of* Åstrand *and* Rhyming

These results have clarified the position with regard to the premises outlined in the Section 1. Two of the premises are fully substantiated. They are that, for a task performed against gravity, *O* can be predicted from *W* with a variance that is of the same order as the day-to-day variability in direct measurement of *O* ; that the variation of an individual maximum *H* from the mean for a population is small, so that for young healthy Bantu males a mean maximum *H* of 180 beats/min can be employed in any test procedure without introducing significant errors.

This study did, however, show that one of the premises is not strictly valid. A straight line is a good representation of the relationship between *H* and *O* over most of the range of work rates, but at levels of *W* near the individual maximum the relationship departs sharply from linearity, so that the measured maximum *O* or the asymptotic value of *O* is generally higher than would correspond to the maximum *H* value when the straight line fitted to plots of *H*, *O* is extrapolated to that value. The extent of the ' bias ' is, however, small.

The most important shortcoming, brought out by this study, in Åstrand and Rhyming's proposed procedure is that the straight lines fitted to the plots of four pairs of *H* and *O* values for each of the 26 men DO NOT have a common point. The difference between resting *O*'s in the 26 men were not significant, but the differences between the *H*'s were very significant. Åstrand and Rhyming's procedure is to draw a straight line between the common point, viz., 60 beats/min at zero *O*, and the measured *H* plotted against estimated *O* for a single rate of stepping. Their method has therefore a number of possible sources of error. The first is the ' bias ' of the common point compared with individual's measured *H* at rest ; the second is the variances of *H* at rest ; the third is the random error in measurement of *H* during stepping ; and the fourth is the ' bias ' of the true *O* during stepping compared with that estimated from the population mean. Only the last two of these sources of error affect the estimate as proposed in this paper. The cumulative effect of all errors on the precision of the estimate of *O* is examined for both methods in the succeeding sections.

4.2. *Alternative method for estimating the individual H/O straight line*

The unsatisfactory features of the Åstrand and Rhyming method may be eliminated firstly, by omitting H's which are not actually measured and, secondly by using additional work rates in order to obtain more data for fixing an individual line. Our proposal is as follows: select four rates of work and determine the subject's H at each of them, estimate his four O's from the average O, W line, fit a straight line through the four O, H points, extrapolate it to $H = 180$, and find corresponding O as an estimate of maximum O.

It is important to make an assessment of the random error associated with the final maximum O determined in this way. There are various sources of error whose effects are, generally, cumulative, and they are:

(i) random observer errors in measuring H;
(ii) day-to-day physiological variations in H and O at given W;
(iii) the error introduced by estimating O from the 'standard' graph of O, W instead of measuring O directly;
(iv) the use of a single value (approximately 180) as maximum H for all subjects.

Many estimates of the combined effect of (i) and (ii) are given in this paper (e.g. Tables 9, 6 and 11) and, as a representative value for specimen calculations of var $(O_{\max})$ the figure var $(H) = 40$ was used.

Table 11. Pooled residual variances for each subject

	Heart rate				
No.	X^2	d.f.	P	Pooled Res. Var.	Asymptote
1	10·7	9	0·30	33·36	178·9
2	4·4	7	>0·70	43·32	180·2
3	20·8	10	<0·025 >0·01	44·05	171·1
4	22·5	7	<0·01	34·74	183·0
4	13·0	5	<0·025 >0·01	107·45	190·63
5	21·3	8	<0·01 >0·005	18·65	194·3
6	36·5	7	<0·0005	78·81	190·4
			Oxygen intake		
1	7·8	9	>0·50	0·0164	3·106
2	9·2	7	>0·20	0·0090	2·901
3	15·1	10	>0·10	0·0297	3·108
4	6·8	7	>0·40	0·0267	3.640
4	11·5	7	>0·10	0·0322	3·055
5	18·7	8	<0·025 >0·01	0·0230	3·002
6	12·9	7	<0·1 >0·05	0·0212	2·714

As regards (iii) and (iv) it is important to note the following: if measurements of O and maximum H are actually made an unbiased estimate of maximum O can be made. This means that if the procedure were repeated a large number of times the average of all determinations will approach 'true' maximum O more and more closely as the number of determinations is increased on the same individuals.

When O (and maximum H) are determined from a population average value the situation is quite different. The estimate of maximum O is almost certainly biased, i.e. repeated determinations of maximum O and averaging do not ensure that true maximum O is approached more and more closely. On the other hand, 'probability' limits for maximum O can still be calculated to include true maximum O with given probability. These limits are not confidence limits in the usual sense ; their width does not decrease to zero with increasing number of observations.

Estimated maximum *O* and its standard deviation was calculated by the method outlined in the Appendix for six subjects and the results are given in

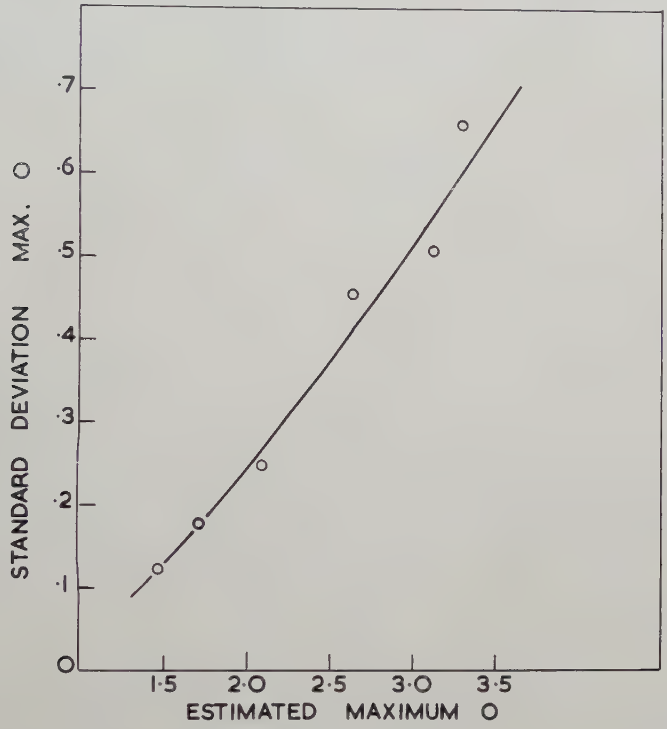


Figure 5. Standard deviations of maximum oxygen intake are given for men with various levels of estimated maximum oxygen intake. A smooth curve has been fitted to the values.

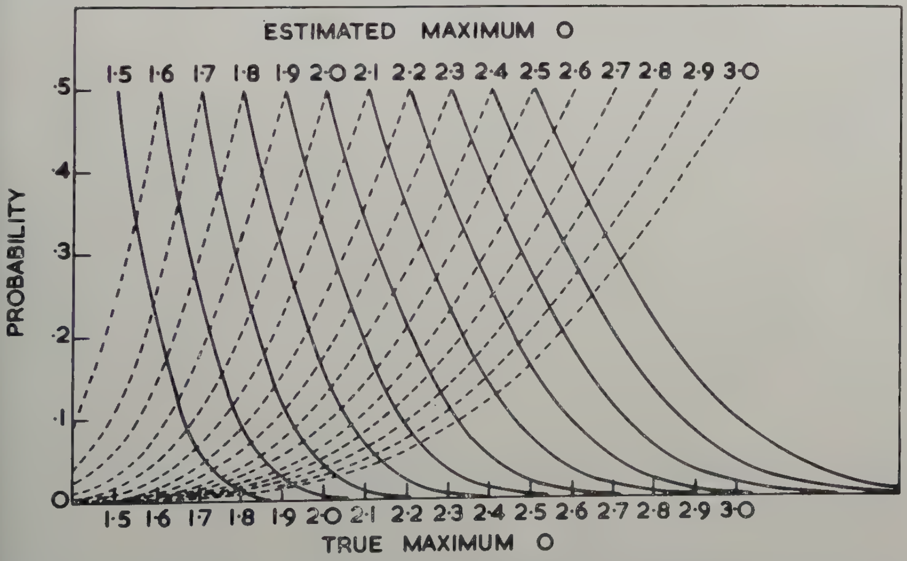


Figure 6. Curves giving the probability of true maximum *O* being less than (dotted lines) or greater than (solid lines) any chosen value for given observed values of estimated maximum *O*.

Table 12. The values of standard deviation of maximum O were plotted against estimated maximum O and it will be seen in Fig. 5 that the points cluster closely about a smooth curve. It was assumed that the standard deviation of maximum O may be represented with sufficient accuracy by a simple mathematical function of maximum O and that the smooth curve drawn in Fig. 5 is, for present purposes, a reasonable approximation to this function.

These results are used in the following routine classification scheme. Using the standard deviations in Table 12 and assuming normality of the distributions involved, the graphs in Fig. 6 were plotted. These are interpreted as follows : the dotted curve decreasing to the left of its starting point (estimated maximum O) on the top scale gives the probability that the value along the bottom scale exceeds the true value of maximum O, while the solid curve decreasing to the right of its starting point gives the probability that the value along the bottom scale is less than the true value of maximum O.

Table 12. Estimated maximum oxygen intake and standard deviations of six untrained subjects

No.	$\mathbf{O}_\mathrm{m}$	s.d.$\mathbf{O}_\mathrm{m}$
10	2·09	0·248
11	1·47	0·122
15	1·71	0·176
21	3·28	0·660
24	2·64	0·456
28	3·12	0·507

Two examples of how these curves may be used are :

(i) Suppose that subjects are to be selected for some fairly arduous tasks for which it is considered unsafe or undesirable to have a maximum O of less than 1·5 l./min. At the same time as large a number of men as possible must be accepted for these tasks. This means that, with the lower limit of 1·5, an upper cut-off O_0 must be chosen so that it is the smallest value giving a reasonably small chance of a man whose true maximum O is 1·5 being selected for the hard tasks. If we want this chance to be 0·05 or less we find that $O_0 = 1{\cdot}9$ would be satisfactory because the probability of true maximum O being less than 1·5, given that the estimated maximum O is 1·9, is 0·034.

(ii) Suppose that two limits 2·0 and 3·0 are decided upon. Those men with estimated maximum $0 \leqslant 2{\cdot}0$ are assigned to light tasks and those with estimated maximum $O \geqslant 3{\cdot}0$ having true maximum $O \leqslant 2{\cdot}0$ is 2·7/100, and if more information about the distribution of estimated maximum O is available the overall risk may be calculated. The ‘ waste ’ if the accent is on high-capacity men is represented by the probability that a man with estimated maximum $O \leqslant 2{\cdot}0$ has true maximum $O \geqslant 3{\cdot}0$ and is 0·001 in this case.

4.3. *Comparison of errors in estimates of maximum O by Åstrand and Rhyming's method and by the present procedure*

One of the main sources of error in estimating maximum O by the method of Åstrand and Rhyming is in using a point of common origin as one of the two plots of H against O to which a straight line is fitted. The study on 26 men brought out the fact that while the O's at rest are not significantly different, the resting H values are very significantly different. Hence using a common point introduces an additional random error. The error in estimating maximum O by Åstrand's method due to the variances ' between subjects ' of resting H values alone was calculated by the method given in Appendix 6.5 on Subject No. 28. Using the variance of H_1 of 120 beats/min, then

Variation of maximum $O = 0{\cdot}53$ s.d. 0·75 (Åstrand and Rhyming method)
$= 0{\cdot}26$ s.d. 0·51 (present method)

In addition, the Åstrand and Rhyming method has a bias due to the common point having been chosen as a heart rate of 60 beats/min at zero oxygen intake instead of using the population mean H and O values. As an example, in a man of 3 l./min maximum oxygen capacity, the bias could be of the order of 0·2–0·4 l./min, depending upon the value of the heart rate observed.

The further point should be made that in the present study the intention has been that the procedures in the practical testing of Native recruits should be carried out by non-scientific personnel. It is, naturally, relatively simple to recognize a gross error in heart rate measurement when four values at different rates of work are plotted against estimated oxygen intake. When only one measurement of H is made at a single rate of step climbing such gross errors cannot be observed.

The method of fitting a straight line to plots of four measured H values and predicted O values was proposed also with a view to maintaining simplicity in the procedure. This proposed method is less accurate than fitting a straight line by least squares. When, however, accuracy greater than the quoted variances of maximum O is required, then fitting a line by least squares is the method of choice.

§ 5. CONCLUSIONS

5.1. Oxygen intake of an individual can be estimated with a coefficient of variation of 6 per cent from a population mean straight line relating oxygen intake for various rates of stepping to rate of work. This value is of the same order of magnitude as the residual variance of oxygen intake due to day-to-day variations and observer errors. The random error in estimating oxygen intake from a ' standard ' oxygen intake/work rate graph is therefore no greater than the variances in measured oxygen intake due to error and physiological variation. Measuring the ventilation volume during work does not improve the accuracy of the estimate of oxygen intake by more than a small percentage and the procedure is not therefore warranted.

5.2. The individual variability of maximum heart rate about the population mean is small. The mean for six trained men is 184·1 beats/min with a standard deviation of 8·4 ; and of untrained men it is 179·3 beats/min with a standard deviation of 5·9. The means are therefore not significantly different. Taking a maximum heart rate of 180 beats/min as the population mean for Bantu introduces a very small error in the determination of maximum oxygen intake.

5.3. Heart rate is a linear function of oxygen intake over most of the range of work up to the individual maximum. Near the maximum, however, there is a sharp departure from linearity in most individuals, so that the maximum oxygen intake actually measured is greater by about 0·3 l./min than that estimated by extrapolating the straight line to the maximum heart rate level and reading off the corresponding oxygen intake value on the graph.

5.4. These conclusions are the basic premises upon which various rapid methods of estimating an individual's maximum oxygen intake are founded.

The detailed procedures differ, but the essence of the methods is similar. The procedure is that heart rate is measured at one or more rates of work; oxygen intake is estimated from a ‘ standard ’ graph relating oxygen intake to rate of work; pairs of heart rate and oxygen intake values for various rates of work are plotted on a heart rate/oxygen intake ‘ grid ’; a straight line is fitted and extrapolated to the maximum heart rate value and the corresponding oxygen intake value is an estimate of the individual’s maximum oxygen intake. The procedure proposed by Åstrand and Rhyming (1954) is based on a single measurement of heart rate at ONE rate of work. It has four possible sources of error. The heart rate value at rest may be biased; it has a random error; measurements of heart rate during work also have a random error; estimating oxygen intake from a ‘ standard ’ graph has a random error. Some of these errors are eliminated by a procedure proposed by us which is based on fitting a straight line to the plot of four pairs of heart rate and oxygen intake values. The errors of these two methods have been compared. The standard deviation of maximum oxygen intake of a test subject using Åstrand and Rhyming’s method was 0·73, compared with 0·51 standard deviation of our method. In addition, Åstrand and Rhyming’s method has a systematic bias of about 0·3 l. oxygen intake/min. The accuracy of estimation by the present method could be improved by fitting the straight lines by least squares.

Les hypothèses sur lesquelles sont basées des méthodes simples d’évaluation de la consommation maximale d’oxygène chez un individu ont été éprouvées expérimentalement. Ces hypothèses sont :

(*a*) La fréquence cardiaque et la consommation d’oxygène sont réciproquement des fonctions linéaires pour toutes les puissances jusqu’à la puissance maxima d’un individu.

(*b*) La consommation d’oxygène d’un individu diffère très peu de la droite moyenne exprimant, à l’intérieur de la population, la relation entre la consommation d’oxygène et le travail de telle sorte que la consommation d’oxygène pour une épreuve de travail effectuée contre la pesanteur peut être évaluée avec une bonne précision à partir du travail effectué.

(*c*) La variabilité individuelle de la fréquence cardiaque maxima de part et d’autre de la moyenne de la population est suffisamment petite pour autoriser l’emploi de cette moyenne dans une épreuve d’effort courante sans introduire de grandes erreurs.

Les hypothèses (*b*) et (*c*) sont pleinement vérifiées par les résultats exposés ici. L’hypothèse (*a*) n’est pas strictement validée, du fait d’un écart de l’ordre de o, 32 l/mn d’oxygène par rapport à la relation linéaire entre la fréquence cardiaque et la consommation d’oxygène pour des niveaux de travail élevés, chez la plupart des individus.

Les erreurs imputables à la méthode proposée par Åstrand (1954) basée sur ces hypothèses ont été examinées et comparées avec celles de l’autre version de cette méthode, décrite dans le présent article. La variance de la consommation maximale d’oxygène par la méthode d’Åstrand, due aux erreurs de mesure de la fréquence cardiaque seule, est de l’ordre de 0,53, comparée à une variance de 0,23 par le procédé décrit dans cet article.

De plus une mesure unique de la fréquence cardiaque par la méthode d’Åstrand introduit une erreur d’environ 0,3 l d’oxygène/minute dans l’évaluation d’une consommation d’oxygène maxima aérobie de 3,0 l/mn. La variance de la consommation maxima d’oxygène prédite à l’aide de la présente méthode peut être réduite en utilisant la droite des moindres carrés ajustée aux quatre points expérimentaux correspondant aux fréquences cardiaques et aux consommations d’oxygène observées pour quatre niveaux de travail différents.

Die Voraussetzungen, auf die sich einfache Methoden oder Schätzungen der maximalen Sauerstoff-Aufnahme eines Individuums stützen, wurden geprüft. Diese Voraussetzungen sind folgende :

(*a*) Herzfrequenz und Sauerstoff-Aufnahme sind einander linear proportional durch das gesamte Arbeitsbereich bis zu den Maximalwerten eines Individuums.

(*b*) Die Sauerstoff-Aufnahme des Einzelindividuums weicht sehr wenig von der Hauptgeraden ab, welche die Sauerstoff-Aufnahme und Arbeitsschwere des Durchschnitts der Bevölkerung zueinander in Beziehung setzt, sodass die Sauerstoff-Aufnahme für eine Arbeit welche gegen die Schwerkraft geleistet wird, mit genügender Genauigkeit aus der Arbeitsschwere geschätzt werden kann.

(*c*) Die individuelle Streuung der maximalen Herzfrequenz um den Durchschnittswert der Bevölkerung ist genügend klein, um diesen Durchschnittswert in einem Routine-Test-Verfahren ohne grossen Fehler gebrauchen zu können.

Die Angaben (*b*) und (*c*) sind hinreichend durch die hier angeführten Ergebnisse begründet. Die Angabe (*a*) ist nicht streng gültig, da es bei hoher Arbeitsschwere bei den meisten Individuen eine Neigung der gradlinigen Beziehung von Herzfrequenz und Sauerstoff-Aufnahme in der Grössenordnung von 0,3 Liter Sauerstoff/Min. gibt.

Die Fehler, die der auf der gleichen Voraussetzung beruhenden Methode Åstrand's (1954) anhaften, wurden untersucht und mit den Fehlern eines anderen Verfahrens verglichen, das in der vorliegenden Arbeit beschrieben wird. Die Streuung der maximalen Sauerstoff-Aufnahme bei Anwendung der Åstrand-Methode, die allein aus Messfehlern der Herzfrequenz entstehen, liegt in der Grössenordnung von 0,53 verglichen mit 0,26 bei dem hier beschriebenen Verfahren. Zudem führt die alleinige Herzfrequenz-Messung der Åstrand'schen Methode zur Ueberschätzung einer individuellen maximalen Sauerstoff-Aufnahme von 3,0 l/min um ungefähr 0,3 l/min. Die Streuung der maximalen Sauerstoff-Aufnahme-Messung durch die dargestellte Methode kann dadurch verkleinert werden, dass man mit der Methode der kleinsten Quadrate eine gerade Linie durch 4 Paare von Herzfrequenz- und Sauerstoff-Aufnahme-Messungen für 4 Grössen der Arbeitsschwere legt

REFERENCES

CHRISTENSEN, E. H., 1953, Physiological valuation of work in Nykroppa Iron Works. *Symposium on Fatigue* (Ed. W. F. Floyd and A. T. Welford). (London : Lewis) 93.

BRODY, S., 1945, *Bioenergetics and growth* (New York : Rheinhold). 906.

ÅSTRAND, P.-O., 1952, Experimental studies of physical working capacity in relation to sex and age (Stockholm : Munksgaard).

TAYLOR, H. L., BUSKIRK, E. and HENSCHEL, A., 1955, Maximal oxygen intake as an objective measure of cardio-respiratory performance. *J. appl. Physiol.*, **8** 73.

WYNDHAM, C. H., STRYDOM, N. B., MARITZ, J. S., MORRISON, J. F., PETER, J., and POTGIETER, Z. U., 1959, Maximum oxygen intake and maximum heart rate during strenuous work. *J. appl. Physiol.*, **14**, 927.

ÅSTRAND. P.-O., and RHYMING, I., 1954, A nomogram for calculation of aerobic capacity (physical fitness) from pulse rate during submaximal work. *J. appl. Physiol.*, **7**, 218.

STATISTICAL APPENDIX

1. *Variability of H and O at Different Levels of W Step Tests*

The effects of different rates of W on the 'residual variability' of H and O is small and irregular, although there appeared to be a trend in some instances. Bartlett's test was used to test the residual variances for homogeneity and the values of χ^2, the test criterion, are given in Table 9. They indicate that the hypothesis of uniform residual variation is reasonable, therefore the individual residual variances were pooled. The pooled values are also given in Table 9, and it seems that the residual variance of H is characteristic of the individual whereas that of O is not, because testing, again by Bartlett's test, between individuals gives the results

Oxygen intake $\chi_3^2 = 4{\cdot}51$ (not significant)
Heart rate $\chi_3^2 = 15{\cdot}13$ (very significant)

Bicycle Ergometer Tests

Tests of homogeneity of the residual variances of H and O were carried out, and Table 11 contains the results of the test of homogeneity and pooled residual variances for H and O for each of these six individuals. Some of the individual values of χ^2 are high but plotting the individual variances against W showed no clear trend, so that pooling of the residual variances appears to be justifiable.

2. *Estimates of Variances and Covariances of $O|W$ for 26 Men*

Estimates of α_2's and β_2's for the 26 men are given in Table 13, together with their individual variances and covariances. The variances and covariances between α_2's and β_2's, calculated from values in Table 13, are :

$$\begin{array}{llll}
\mathrm{var}\ (\boldsymbol{\alpha}_2) & = & 0{\cdot}004759 & \\
\mathrm{var}\ (\boldsymbol{\beta}_2) & = & 0{\cdot}0^{6}1040/36 & 0{\cdot}0^{8}2829 \\
\mathrm{cov}\ (\boldsymbol{\alpha}_2\beta_2) & = & -0{\cdot}0^{4}1168/6 & -0{\cdot}0^{5}195
\end{array}$$

Using these numbers, estimates of var (α_2), etc., may be found because it is approximately true that

$$\mathrm{var}\ (\alpha_2)+\sum\{\mathrm{var}(\boldsymbol{\alpha}_2|\alpha_2)\} = \mathrm{var}\ (\boldsymbol{\alpha}_2),\ \text{etc.}$$

From the Table :

$$\begin{array}{llll}
\text{mean var } (\boldsymbol{\alpha}_2|\alpha_2) & = & 0{\cdot}005415 & \\
\text{mean var } (\boldsymbol{\beta}_2|\beta_2) & = & 0{\cdot}0^{7}5187/36 & 0{\cdot}0^{8}1441 \\
\text{mean cov } (\boldsymbol{\alpha}_2\beta_2|\alpha_2\beta_2) & = & -0{\cdot}0^{4}1239/6 & -0{\cdot}0^{5}206
\end{array}$$

and finally

$$\begin{array}{ll}
\mathrm{var}\ (\alpha_2) & =0{\cdot}0^{3}244 \\
\mathrm{var}\ (\beta_2) & =0{\cdot}0^{7}5213/36 \\
\mathrm{cov}\ (\alpha_2\beta_2) & =0{\cdot}0^{6}71/6
\end{array}$$

Table 13. Estimates of parameters of straight line fitted to O, W and variances and covariances

Subject No.	α	β	var (α)	var (β)	$-$cov (α, β)
7	0·309	0·00255	0·00191	$0{\cdot}0^{7}229$	$0\ 0^{5}505$
8	0·284	0·00255	0·00135	$0{\cdot}0^{7}190$	$0{\cdot}0^{5}388$
9	0·417	0·00185	0·00483	$0{\cdot}0^{7}682$	$0{\cdot}0^{4}139$
10	0·371	0·00194	0·00448	$0{\cdot}0^{7}613$	$0{\cdot}0^{4}127$
11	0·301	0·00268	0·00507	$0{\cdot}0^{7}759$	$0{\cdot}0^{4}150$
12	0·247	0·00276	0·00219	$0{\cdot}0^{7}314$	$0{\cdot}0^{5}634$
13	0·401	0·00186	0·00143	$0{\cdot}0^{7}134$	$0{\cdot}0^{5}335$
14	0·305	0·00264	0·00520	$0{\cdot}0^{7}810$	$0{\cdot}0^{4}157$
15	0·258	0·00234	0·00293	$0{\cdot}0^{7}458$	$0{\cdot}0^{5}886$
16	0·463	0·00179	0·01636	$0{\cdot}0^{6}188$	$0{\cdot}0^{4}424$
17	0·320	0·00275	0 00678	$0\ 0^{7}700$	$0{\cdot}0^{4}167$
18	0·436	0·00229	0·01345	$0{\cdot}0^{6}192$	$0{\cdot}0^{4}389$
19	0·282	0·00234	0·00014	$0{\cdot}0^{8}183$	$0{\cdot}0^{6}465$
20	0·375	0·00255	0·00443	$0{\cdot}0^{7}493$	$0{\cdot}0^{4}113$
21	0·287	0·00251	0·00214	$0{\cdot}0^{7}350$	$0{\cdot}0^{5}663$
22	0·287	0·00243	0·00020	$0{\cdot}0^{8}309$	$0{\cdot}0^{6}598$
23	0·323	0·00271	0·00007	$0{\cdot}0^{9}837$	$0{\cdot}0^{6}188$
24	0·240	0·00227	0·00406	$0{\cdot}0^{7}589$	$0{\cdot}0^{4}118$
25	0·373	0·00173	0·01198	$0{\cdot}0^{6}181$	$0{\cdot}0^{4}356$
26	0·330	0·00230	0·00139	$0{\cdot}0^{7}186$	$0{\cdot}0^{5}389$
27	0·536	0·00224	0·00894	$0{\cdot}0^{7}750$	$0{\cdot}0^{4}198$
28	0·330	0·00227	0·00456	$0{\cdot}0^{7}527$	$0{\cdot}0^{4}118$
29	0·323	0·00208	0·00386	$0{\cdot}0^{7}531$	$0{\cdot}0^{4}110$
30	0·336	0·00201	0·00640	$0{\cdot}0^{7}797$	$0{\cdot}0^{4}173$
31	0·350	0·00185	0·00257	$0{\cdot}0^{7}353$	$0{\cdot}0^{5}729$
32	0·309	0·00250	0·00065	$0{\cdot}0^{8}832$	$0{\cdot}0^{5}178$

This means that at a W of 900 ft lbs/min for example, the variance of true O is $0{\cdot}0^{3}244+(900)^2\times 0{\cdot}0^{7}5213/36+1800\times 0{\cdot}0^{6}71/6 = 0{\cdot}001630$. This value is smaller than the residual variances of O due to day-to-day variations.

3. *The Effect of Measuring V (Respiratory Ventilation) on the Precision of Estimating O from W*

The effect of V on precision of estimating O from W was examined as follows :

For given W, true O and true V or true O and V (estimated by using measurements of V) are, from man to man, random variables, and unless they are independent, it may be expected that

$$\text{var } (O|W\boldsymbol{V})|\text{var } (O \leqslant W)$$

so that making measurements of V should enable one to set narrow limits to the individual maximum oxygen intake.

Values of α_3 and β_3 in the $V = \alpha_3 + \beta_3 W$ relationship were calculated and, using var $(\alpha_3|\alpha_3)$ and var (α_3), etc., as above, the correlation between O and V was estimated. For W's of 400 ft lb/min and above this correlation was about 0·4 and, assuming that the order of the reduction in var $(O|W, \boldsymbol{V})$ below var $(O|W)$ is similar to the occurrence in normal distribution theory (where var $(O|W, {}_1) =$ var $(O|W)$ $(1–r^2)$ and r is the correlation coefficient) s.d. $(O|W, V) = 0{\cdot}9$ s.d. $(O|W)$.

4. *Calculation of 'Probability' Limits of Maximum O*

For calculation of the limits, the O's (and H_m) are most conveniently regarded as random variables with variances calculated (in the case of O) using var (α_2), var (β_2), cov $(\alpha_2\beta_2)$. The observed variance of H_m between the six subjects (Table 9) is used.

The actual procedure selected for determining $O_{\max}$ from the four pairs $(o_1, h_1) \ldots (o_4, h_4)$ is

$$\boldsymbol{O}_m = \boldsymbol{\gamma} + \boldsymbol{\gamma} H_m$$

where

$$\boldsymbol{\delta} = \frac{(O_1 + O_2) - (O_3 + O_4)}{(h_1 + h_1) - (h_2 + h_4)}$$

$\boldsymbol{\gamma} = \overline{O} - \boldsymbol{\delta}\bar{h}$, O and $\bar{h}$ being the means of the four O's and h's respectively. This choice of $\boldsymbol{O}_m$ was dictated largely by considerations of simplicity because it is anticipated that relatively unskilled personnel will be using the method on mines.

The variance of $\boldsymbol{O}_m$ is given approximately by

$$\text{var } (\boldsymbol{O}_m) = u'mu$$

where $u' = (1\ H_m, ')$

$$\text{and } M = \begin{pmatrix} \text{var } (\boldsymbol{\gamma}) & \text{cov } (\boldsymbol{\gamma}, \boldsymbol{\delta}) & \text{cov } (\gamma, H_m) \\ \text{cov } (\boldsymbol{\gamma}, \boldsymbol{\delta}) & \text{var } (\boldsymbol{\delta}) & \text{cov } (\delta, H_m) \\ \text{cov } (\boldsymbol{\gamma}, H_m) & \text{cov } (\boldsymbol{\delta}, H_m) & \text{var } (H_m) \end{pmatrix}$$

The following notation was used in calculating the elements of M

$x_1 = (O_1 + O_2) - (O_3 + O_4)$ $\sum(x_1) = \xi_1$

$x_2 = (h_1 + h_2) - (h_3 + h_4)$ $\sum(x_2) = \xi_2$

${}_1 = \overline{O}, f_2 = \bar{h}$; $\sum(f_1) = \phi_1, \sum(f_2) = \phi_2$

var $(x_1)/\xi_1{}^2 = c_1\,.$, var $(x_2)/\xi_2{}^2 = c_2\,.$, var $(f_1)/\phi_1{}^2 = c\,._1$,
var $(f_2)/\phi_2{}^2 = c\,._2$, cov $(x_1, f_1)/\xi_1\phi_1 = c_{11}$,
cov $(x_2, f_2)/\xi_2\phi_2 = c_{22}$, and the results derived were :—

var $(\boldsymbol{\gamma})$ $= \phi_1{}^2 c\,._1 + \delta\phi^2(c_1\,. + c_2\,. + c\,._2 - 2c_{22}) - 2\delta\phi_1\phi_2 c_{11}$

var $(\boldsymbol{\delta})$ $= \delta^2(c_1\,. + c_2\,.)$

cov $(\boldsymbol{\gamma}, \boldsymbol{\delta})$ $= \phi_1\delta c_{11} - \phi_2\delta^2\ (c_1\,. + c_2\,. - c_{22})$

cov $(\boldsymbol{\gamma}, H_m)$ $=$ cov (f_1, H_m) $-$ (ϕ_2/ξ_2) cov (x_1, H_m)

cov $(\boldsymbol{\delta}, H_m)$ $=$ cov $(x_1, H_0)/\xi_2$.

To get the values of var (x_1), etc., is fairly simple. Writing $W = (W_2 + W_2) - (W_3 + W_4)$ and w for the mean of the Wi's we have

$$\begin{aligned} \text{var } (x_1) &= w^2 \text{var}(b_2) \\ \text{var } (f_1) &= \text{var } (a_2) + \overline{w}^2 \text{ var } (b_2) + 2\overline{w} \text{ cov } (a_2 b_2) \\ \text{cov } (x_1 f_1) &= w \text{ cov } (a_2 b_2) + W\overline{w} \text{ var } (b_2) \end{aligned}$$

$$\begin{aligned} \text{var }(x_2) &= 4\sigma^2_h \\ \text{var }(f_2) &= \sigma^2_h/4 \\ \text{cov }(x_2 f_2) &= O \\ \text{cov }(x_1 H_0) &= W \text{ cov }(H_m b_2) \\ \text{cov }(f_1 H_0) &= \text{cov }(H_m a_2) + \bar{w} \text{ cov }(H_m b_2) \end{aligned}$$

As explained above, observed var (H_m) was used, and the same applied to cov (H_m, a_2) and cov (H_m, b_2). These estimates are very rough but it was, in fact, established that they do not affect the result very much. The main contribution to var ($\boldsymbol{O}_m$) is due to var (H).

5. *Error in Estimating Maximum O by Proposed Procedures*

$$\text{Variance maximum } O = \frac{(180-H_2)}{(H_2-H_1)^2}\,(O_2-O_1)^2 \text{ var } H_1$$

where H_1 and O_1 are the assumed common point, viz., resting values, and H and O_2 are H and O values at work.

The 'between subject' variance in resting H can be determined from residual variances on the 26 men and the four trained subjects, and is

$$(4^2)(10) = 40 = 120$$

Hence, taking hypothetical values of

$$\begin{aligned} H_1 &= 80 \text{ and } O_1 = 0{\cdot}3 \text{ l./min} \\ H_2 &= 140 \text{ and } O_2 = 2{\cdot}1 \text{ l./min} \end{aligned}$$

Then

$$\begin{aligned} \text{maximum } O &= 3{\cdot}3 \text{ l./min} \\ \text{s.d.} &= 0{\cdot}25 \end{aligned}$$

$$\text{and coefficient of variation} = \frac{0{\cdot}25 \times 100}{3{\cdot}3} = 8 \text{ per cent}$$

This figure is of the same order as the s.d. of maximum O given in a preceding section. The final s.d. from two observations may, however, be expected to be larger than that from four observations, and can be calculated as follows :

$$\begin{aligned} \text{Var }(\boldsymbol{O}_m) = \text{var }(O_2)&\left\{1+\frac{180\ 9\ H_2}{(H_2-H_1)}\right\}^2 + \left\{\frac{(180-H_2)^2}{(H_2-H_1)}\right\} \text{var }(O_1) \\ &-2\left\{1+\frac{180-H_2}{H_2-H_1}\right\}\left\{\frac{180-H_2}{H_2-H_1}\right\} \text{cov } O_1O_2 \\ &+\left\{-\frac{(H_2-H_1)-(180-H_2)}{(H_2-H_1)}\right\} \text{var }(H_2)\,(O_2-O_1)^2 \\ &+\left\{\frac{(180-H_2\,(O_2-O_1))}{(H_2-H_1)^2}\right\}^2 \text{var }(H_1) \end{aligned}$$

where H_1 is the assumed common resting H, and

$$\begin{aligned} O_2 &= A+BW_2 \\ O_1 &= A+BW \\ \text{var } O_1 &= \text{var }(a)+2W_1 \text{ cov }(ab)+W_1^2 \text{ var }(b) \\ \text{var } O_2 &= \text{var }(a)+2W_2 \text{ cov }(ab)+W_2^2 \text{ var }(b) \\ \text{cov }(O_2O_1) &= \text{var }(a)+(W_1+W_2) \text{ cov }(ab)+W_1W_2 \text{ var }(b) \end{aligned}$$

as before.

Applying this procedure to one of the 26 subjects, No. 28, and using the variance of H_2 of 120 beats/min, we then have

$$\begin{aligned} \text{var } O_m &= 0{\cdot}53, \text{ s.d. } 0{\cdot}73 \text{ (Åstrand and Rhyming method)} \\ &= 0{\cdot}26, \text{ s.d. } 0{\cdot}51 \text{ (present method)} \end{aligned}$$

THE ESTIMATION OF BODY MEASUREMENTS OF BRITISH POPULATION IN RELATION TO SEAT DESIGN

By D. Barkla

Furniture Development Council, Research Department, 2 Dalmeny Avenue, London, N.7.

Reports of surveys of human measurements are not always easily accessible, and their information about dimensions relevant to seat design is often mixed up with other matter. This paper brings together the principal published information on relevant dimensions. Most of the populations surveyed are not themselves important to British designers, but estimates of the measurements of young British adults are derived from the published material.

§ 1. Introduction

There are many different opinions about how chairs should be designed. One thing, however, seems to be generally agreed: that the dimensions of the chair should be in reasonable agreement with the dimensions of the person sitting in it. It is also widely held that he should be able to change position easily. Good fit alone may not ensure comfort, but bad fit will almost certainly ensure discomfort.

Unfortunately, human body measurements that designers could use have often been contained in surveys and reports which are not easily accessible. This paper summarizes a number of these. Few were British or were intended to provide data for designers of seats, but in this paper estimates of British measurements have been derived from them. The assumptions underlying the derivation are pointed out. Nearly all the data, and hence the estimates, refer to adults under the age of 40.

§ 2. Available Measurements of Selected Dimensions

The published dimensions considered important to seat designers are set out in Table 1. They are usually defined as follows:

Subjects are measured naked, and, except for stature and shoulder width, are seated on a horizontal table with trunk erect and looking directly forward, upper arms vertical and forearms horizontal, underpart of thighs just touching the table, knees together, legs vertical, and feet on a horizontal rest.

Stature: subjects standing erect looking directly forward; from floor to vertex, hair pressed down.

Shoulder width: standing erect with arms at sides and palms forward; distance between outer aspects of shoulders, across the middle deltoid muscles.

Sitting height: from seat to vertex, hair pressed down.

Shoulder–seat: from seat to top of shoulder (highest bony point) or from seat to top of breast-bone (sternum). The two dimensions have practically dentical values.

Lumbar height: from seat to maximum inward deflection of the back.

Elbow–seat: from from seat to bottom of elbow.

Buttock–back of calf: from back of buttocks to back of calf at its maximum diameter

Table 1. Means and standard deviations of sel

Source number	Sex	Nationality	Number in sample		Age		Stature		Weight		Sitting height		Shoulder–seat		Wa hei
			Sub-sample	Total sample	μ or m	σ or w	μ	σ	μ	σ	μ	σ	μ	σ	μ
1 and 2	Male	American		2960			69·2	2·43	153·1	17 58	36·4	1·20	23·8	0·93	
1	,,	,,		584			67·9	(2·45)	147·1	(15·81)	35·9	(1·19)	23·3	(0·96)	
2	,,	,,		103	μ34·1	σ 5·98	69·0	2·28	167·1	27·04	36·6	1·06	23·7	1·06	
2	,,	,,		271	μ38·9	σ 8·13	68·3	2·41	166·9	27·93	36·2	1·16	23·7	1·00	
3	,,	,,		25000	m22	w15–40	68·5	2·51	154·8	20·49	35·8	1·34	23·0	1·17	
5	,,	,,		4063	μ27·9	σ 4·22	69·1	2·44	163·7	20·86	35·9	1·29	23·3	1·14	9·2
5	,,	,,	1184		μ28·9	σ 3·61	69·4	2·3	165·7	20·5	36·1	1·24	23·4	1·09	9·3
5	,,	,,	774		μ26·9	σ 2·14	69·5	2·22	164·0	20·51	36·1	1·27	23·4		9·3
5	,,	,,	398		μ32·8	σ 2·52	69·4	2·46	167·8	20·31	36·1	1·26	23·4		9·3
5	,,	,,	81		μ36·5	σ 3·33	69·1	2·20	168·4	20·57	36·0	1·15	23·5		9·2
5	,,	,,		94	μ32·2	σ 3·17	69·6	2·46	173·2	19·77	36·2	1·36			
6	,,	,,		1959	μ38·5	w17–89	69·0	(2·89)	166·7	(25·1)					
7	,,	Swedish		295			69·3								
7	,,	,,		274											
7	,,	,,		248											
8 and 9	,,	British		2241	μ21·2	w18–38	68·4	2·47	139·6	16·09	35·7	1·26			
9	,,	,,		529	μ23·2	σ 3·26	68·5	2·37	144·9	15·90	36·2	1·24			
10	,,	,,		361	m24		71·3	1·50			37·7	1·01			
10	,,	,,		1844	m24		67·9	2·39			36·3	1·17			
13	,,	,,		various	μ23·0	σ 1·52	68·1	2·29	148·3	16·85	35·4	1·33	21·6	1·06	
1	Female	American		447			64·9	(1·75)	128·6	(12·4)	34·1	(1·02)			
1	,,	,,		152			63·5	(2·33)	121·9	(12·4)	33·7	(1·19)			
4	,,	,,		8500	m25	w18–50	63·9	2·41	131·9	19·96			21·2	1·09	
11	,,	,,		10042	m32		63·2	2·48	133·5	25·98					
6	,,	,,		1908	μ35·5	w17–84	64·9	(2·51)	133·1	(21·5)					
7	,,	Swedish		363			64·2								
7	,,	,,		251											
7	,,	,,		59											
12	,,	British	2366			w18–29	63·4	2·48	126·6	18·37					11·4
12	,,	,,	1554			w30–44	63·1	2·63	134·6	22·56					11·7
12	,,	,,	1075			w45–64	62·1	2·55	143·8	25·56					12·2
12	,,	,,		4995		w18-64	63·0	2·59	132·8	22·45					11·7

K

Weight is tabulated in pounds, other dimensions in inches. μ=mean or median (dimen
m=median (age of subjects), σ=standard deviation, w=range. (Values bracketed in Ta
have been estimated for this paper.)

The following notes refer to each numbered source in Table 1, and show as f
possible how samples were constituted, and whether the definitions of dimensions differed
those given above.

1. Randall *et al.* (1946). Samples: males between 60 in. and 76 in. tall, and weighin
tween 120 lb and 200 lb (2960 Aviation Cadets, 584 gunners). Females betwee
specified limits of height and weight (447 pilots, 152 Air Force nurses).
Measurements: subjects sat with legs dangling, not on a foot-rest. Elbow
measured with arms at sides of thighs. Knee height from base of heel to top of knee.
2. MacFarland *et al.* (1953). Samples: 103 champion lorry drivers, 271 regular lorry d
mostly from Texas or New England. Measurements: as for source 1 (but subjects clot
3. Newman and White (1951), Randall and Baer (1951). Sample: about 25 000
U.S. Army men.
4. Randall and Munro (1949) and Randall (1947). Sample: about 8500 U.S. Army w
Measurements: subjects sat with feet on a non-adjustable rest. Knee height from foo
to patella, not to top of knee.
5. Fry and Churchill (1956) and Hertzberg *et al.* (1954). Samples: (i) all available
personnel at 14 Air Force base (4063 men); (ii) 94 Strategic Air Command pilots.
Sub-groups of sample (i): All bomber pilots: under 30, 774 men; over 30, 398
over 34, 81 men; total, 1184 men.
6. Hooton (1945). Samples: 1959 men and 1908 women volunteers over 17, at E
and Chicago railway stations. Measurements: subjects clothed and shod, sitting
adjustable chair with seat tilted at 7° from horizontal. Elbow–seat: ‘ from arm-r
seat ’ (they were not parallel). Buttock–back of calf: from back of chair to back o
Underside of thigh–floor: seat height at back of chair, plus corrections for slope of se
compression of thigh when it projected beyond the front edge of the seat, and for w
who said they took shoes off in trains. Seat width: maximum spread across upper
or buttocks.

mensions collected from various sources

–	Elbow–seat		Buttock–knee		Buttock–back of calf		Knee height		Underside of thigh – floor		Shoulder width		Elbow width		Buttock width	
σ	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ
8			23·6	1·06			22·0	0·94			18·0	0·77	16·7	0·99	14·0	0·75
4)			23·1	(1·04)			21·5	(0·96)			17·7	(0·74)	16·4	(1·11)	13·8	(0·82)
6			23·7	1·09			21·7	1·03			18·7	1·05	17·3	1·67	14·7	1·03
0			23·7	1·12			21·6	0·99			18·3	0·97	17·9	1·68	14·5	1·08
1			23·4	1·12			21·6	1·09			17·9	0·99	17·5	1·54	14·0	0·90
9	9·1	1·04	23·6	1·06			21·7	0·99	17·0	0·77	17·9	0·91	17·3	1·42	14·0	0·87
8	9·2	1·02	23·7	1·03			21·8	0·96	17·1	0·74	18·0	0·89	17·5	1·41	14·1	0·87
	9·2		23·7	0·99			21·8		17·1		18·0	0·86	17·3	1·37	14·0	0·87
	9·2		23·8	1·05			21·7		17·0		18·1	0·91	17·8	1·44	14·2	0·87
	9·3		23·7	1·01			21·6		16·9		18·1		17·8	1·49	14·2	0·84
			23·7	1·01											14·8	0·86
	9·6	(0·89)			18·9	(0·96)			19·0	(0·89)	17·6	(0·67)			15·3	(1·11)
									(17·2)	(0·81)						
					(19·3)	(0·95)										
							(21·7)	(1·05)								
			22·8	1·25												
			23·3	1·03							16·8	0·69				
	10·9	1·18	24·4	0·81			22·3	1·03	15·8	0·42	17·0	0·63	18·0	0·77	15·4	0·70
	10·5	0·94	23·4	1·07			20·8	1·12	15·3	0·53	16·5	0·68	17·7	0·83	15·3	0·69
	8·7	1·04	22·3	1·00	17·9	0·90	19·9	0·99			18·5	0·84			13·9	0·84
8)			22·6	(0·96)			20·1	(0·88)			16·1	(0·88)	15·1	(1·17)	15·0	(1·03)
1)			22·4	(0·88)			19·5	(0·58)			15·7	(0·58)	14·9	(0·88)	15·1	(0·89)
3							18·8	0·98								
6																
	9·7	(0·89)			18·2	(1·04)			18·1	(0·89)	15·7	(1·04)			14·6	(1·04)
									(15·7)	(0·81)						
					(18·7)	(0·97)										
							(20·5)	(0·81)								

Åkerblom (1948). Sample (or samples?): ' non-selected ' Swedish factory and office
rkers. Measurements: subjects clothed but without shoes, sitting, presumably on their
n chairs. Buttock–back of calf: from back of chair to back of calf.

Morant (1943) and Morant and Gilson (1945). Sample: volunteers for R.A.F. flying
ies, mostly from southern England. Men less than 60 in. tall excluded. Measurements:
jects sitting on a form with buttocks and shoulders against a wall.

Morant and Gilson (1945). Sample: R.A.F. aircrew, including 185 pilots (not less than
in. tall) and 14 gunners (not more than 72 in. tall). Measurements: subjects sitting on a
ol adjusted so that tops of thighs were horizontal, and with backs against a measuring
rd. Shoulder width: with arms folded.

Hill and Morant (1945). Samples: Royal Armoured Corps tank personnel: 361 Guards
t less than 69 in. tall); 1844 from other units. Measurements: subjects clothed but
hout boots or tunics, sitting with backs against a measuring board. Elbow–seat: with
s pressed back against board. Underside of thigh–floor: the height of the seat. Elbow
th: distance between outer surfaces of forearms, about 2 in. in front of elbows. Seat
th: distance between outer surfaces of thighs just below greater trochanters, with knees
ether.

O'Brien and Shelton (1941). Sample: volunteers from various women's organizations;
sewives and country women inadequately represented. Measurements: Shoulder–elbow:
h hand on hip, from point of elbow to intersection of armscye and shoulder.

Kemsley (1957). Sample: volunteers from firms and various women's organizations;
king women and Londoners over-represented. Sub-groups: age groups 18–29, 30–44,
45–64. Measurements: subjects wearing their usual foundation garments.

Geoghegan *et al.* (1957) and Roberts (1957). Sample: non-commissioned naval and
al Marines personnel at Portsmouth. Measurements: stature and weight, measured
ctly. Subjects then photographed and measurements later read from various sub-
ples (containing between 61 and 287) of the photographic records. Sitting height was
sured directly on another sample of 63. Buttock–knee and buttock–back of calf: with
h axis horizontal, measured from lumbar plane, not back of buttock. Shoulder width:
h arms hanging. Knee height: from floor to upper margin of patella. (Note also thigh
dth, maximum: $\mu = 15{\cdot}9$ in., $\sigma = 1{\cdot}10$ in.).

Table 2. Body dimensions as proportions of stature : Males

Source	1	1	2	2	3	5	6	7	8	9	10	10	13
Stature (actual)	69·2	67·9	69·0	68·3	68·5	69·1	69·0	69·3	68·4	68·5	71·3	67·9	68·1
Sitting height	0·526	0·529	0·530	0·530	0·522	0·520			0·522	0·528	0·529	0·534	0·520
Shoulder – seat height	0·344	0·343	0·343	0·347	0·336	0·337							0·317
Waist height						0·133	Subjects wearing shoes						
Shoulder – elbow	0·212	0·214	0·213	0·215	0·209	0·207							
Elbow – seat						0·132					0·153	0·155	0·128
Buttock – knee	0·341	0·340	0·343	0·347	0·341	0·341			0·333	0·340	0·342	0·344	0·328
Buttock – back of calf								0·279					0·263
Knee height	0·318	0·317	0·314	0·316	0·315	0·314		0·313			0·313	0·306	0·292
Underside of thigh – floor						0·246		0·248			0·222	0·225	
Shoulder width	0·260	0·261	0·271	0·268	0·261	0·259				0·245	0·238	0·243	0·272
Elbow width	0·241	0·242	0·251	0·262	0·255	0·250					0·252	0·261	
Buttock width	0·202	0·203	0·213	0·212	0·204	0·203					0·216	0·225	0·204

Table 2 (cont.). Body dimensions as proportions of stature: Females

Source	1	1	4	11	6	7
Stature (actual)	64·9	63·5	63 9	63·2	64·9	64·2
Sitting height	0·525	0·530				
Trunk height			0·332		Most subjects wearing shoes	
Shoulder – elbow	0·211	0·214	0·205	0·210		
Elbow – seat						
Buttock – knee	0·348	0·352				
Buttock – back of calf						0·291
Knee height	0·310	0·307	0·294			0·319
Underside of thigh – floor						0·245
Shoulder width	0·248	0·247				
Elbow width	0·233	0·235				
Buttock width	0·231	0·238				

Underside of thigh–floor: from foot-rest to underside of thigh just behind the knee.

Elbow width: distance between outer aspects of elbows.

Buttock width: distance between bony protuberances at upper ends of thighs (the greater trochanters).

Shoulder–elbow: from top of shoulder (highest bony point) to bottom of elbow.

Buttock–knee: from back of buttock to foremost point on patella.

Knee height: from foot-rest to top of muscle mass near lower end of femur.

Waist height: from seat to greatest lateral abdominal indentation, or where the belt is worn.

§ 3. ESTIMATING MEASUREMENTS OF THE BRITISH POPULATION

Table 1 contains very little about people over 40 years old. Older people tend to be shorter, but that cannot be described quantitatively. The estimates given in this section refer to people between 18 and 40 years of age. Within these limits it is possible to use Table 1 to provide an estimate of the means and standard deviations of British measurements for each dimension. The main difficulties are that most of the available measurements are American, that most of the measurements are of service men and women and not, therefore, representative of the whole population, and that the definitions of some dimensions differ from one source to another.

(*a*) *Estimates of means for the male population*

Table 2 shows mean measurements expressed as proportions of stature. In the evidence for males, sitting height, buttock–knee length, and knee height, the only really comparable measurements, show good agreement between British and American values. Several other measurements look to be in reasonable accord in view of stated differences in measuring technique. It is assumed, therefore, that national differences in mean bodily *proportions* may be neglected here.

Table 3. 'Best' estimates of body dimensions as proportions of stature (rough averages from Table 2, allowing for differences of technique: see Key to Table 1)

	Males	Females
Sitting height	0·525	0·525
Shoulder – seat	0·340	0·335
Waist height	0·135	
Shoulder – elbow	0·210	0·210
Elbow – seat	0·130	0·125
Buttock – knee	0·342	0·350
Buttock – back of calf	0·280	0·290
Knee height	0·315	0·310
Underside of thigh – floor	0·245	0·245
Shoulder width	0·260	0·250
Elbow width	0·255	0·250
Buttock width	0·205	0·235

Although very small men are excluded from the services, the mean proportions of servicemen may be used for the whole population, since Table 2 indicates that moderate differences in stature do not lead to substantial differences in proportion (*vide* source 10, and, less strikingly, source 1).

If there is a reliable measure of at least one dimension in the whole British population, proportions from other samples can be used to provide absolute values for the British population. Morant (1947) states that the mean stature for British adult males is 67·5 in. Clements and Pickett (1957) say it is 67·05 in. Both those values raise problems, since they differ sharply from the values of 68 in. or more obtained from servicemen. Perhaps servicemen are much more highly selected for stature than they are supposed to be; or perhaps military training increases apparent stature, by altering posture. If training does increase apparent stature, it may be inappropriate to ‘ scale down ’ all other servicemen’s measurements to suit civilian stature; training seems unlikely to alter all measurements in the same way and to the same degree. However, from the data it can only be assumed that the civilian mean stature is really between 67·05 in. and 67·5 in.; say 67·25 in.

The estimates of means of other dimensions for British adult males are given in Table 4. They are obtained by multiplying 67·25 in. by the ‘ best ’ estimate of each dimension as a proportion of stature (Table 3).

Table 4. Estimated measurements, in inches, of the British population between 18 and 40

Dimension	Men			Women		
	μ	σ	90 per cent range	μ	σ	90 per cent range
Stature	67·25	2·6	63·00–71·50	63·25	2·6	59·00–67·50
Sitting height	35·25	1·4	33·00–37·50	33·25	1·4	31·00–35·50
Shoulder – seat	23·00	1·2	21·00–25·00	21·25	1·2	19·25–23·25
Lumbar height			(8–12)*			
Elbow – seat	8·75	1·1	7·00–10·50	8·00	1·1	6·25– 9·75
Buttock – back of calf	18·75	1·0	17·00–20·50	18·25	1·0	16·50–20·00
Underside of thigh – floor	16·50	0·8	15·25–17·75	15·50	0·8	14·25–16·75
Shoulder width	17·50	1·0	15·75–19·25	15·75	1·0	14·00–17·50
Elbow width	17·25	1·6	14·75–19·75	15·75	1·6	13·25–18·25
Buttock width	13·75	0·9	12·25–15·25	14·75	0·9	13·25–16·25

μ=mean (to nearest 0·25 in.), σ=standard deviation (to nearest 0·1 in.), 90 per cent range=5th and 95th centiles (to nearest 0·25 in.).

* Darcus and Weddell (1947): a sample of sailors; mean stature not given.

(*b*) *Estimates of means for the female population*

In estimating means for the unselected British adult female population from the available material, the same assumptions have to be made as for males. Kemsley (1957) states that the mean stature for British females aged 18-44 is 63·3 in. This agrees well with a value for American civilian women, but not with values for American service women.

The estimates of means for British adult females are given in Table 4. In arriving at the ‘ best ’ estimates of proportions (Table 3) it is assumed, because the data for females is so scanty, that corresponding sources for males and females are similarly biassed. The assumption is supported by the values in Table 2 for shoulder–elbow, which are higher in source 1 than in source 3 (males) and the corresponding source 4 (females).

(c) *Estimates of standard deviations for the male population.*

The standard deviations given in Table 1 agree fairly well among them-elves, but there are several possible sources of disagreement: the most mportant is variation among the populations (in general, the various services) rom which the samples are drawn. Service groups are so selected as to be nore homogeneous than the parent population –that is, they yield reduced tandard deviations of measurements—and they may also be so selected, or rained, as to yield different mean measurements. Inspection of Table 1 and he notes to it shows that the effects of selection are considerable. The least elected groups (sources 3 and 5) have the greatest standard deviations, and re used to provide the best estimate wherever possible.

To compare the variabilities of samples with diverse means, the co-fficient of variation (standard deviation × 100)/mean is used. There is fair oncordance between the samples except where the group is highly selected r where there are clear differences of measuring technique.

Table 5. Coefficients of variation: males

Source	1 n= 2960	3	5	6	7	10 n= 361	10 n= 1844	13	Clements and Pickett 1957
Stature	3·5	3·7	3·5	Subjects wearing shoes	3·6?	2·1	3·5	3·4	3·9
Sitting height	3·3	3·7	3·6			2·7	3·2	3·8	
Shoulder – seat	3·9	5·1	4·9					4·9	
Shoulder – elbow	4·6	5·7	4·8						
Elbow – seat			11·4	9·3		10·8	9·0	12·0	
Buttock – knee	4·5	4·8	4·5			3·3	4·6	4·5	
Buttock – back of calf				5·1	4·9			5·0	
Knee height	4·3	5·0	4·6		4·8	4·6	5·4	5·0	
Underside of thigh–floor			4·5	Subjects wearing shoes	4·7	2·7	3·5		
Shoulder width	4·3	5·5	5·1	3·8		3·7	4·1	4·5	
Elbow width	5·9	8·8	8·2			4·3	4·7		
Buttock width	5·4	6·4	6·2	7·3		4·5	4·5	6·0	

There are far greater differences in coefficients of variation between mensions than between samples. This might mean that some dimensions ally are more variable than others, but it might also mean merely that some mensions are harder to measure than others. Most of the values in Table 5 e reasonably concordant and reasonably low, but the coefficients of variation elbow–seat height and elbow width should certainly be regarded with spicion.

The estimation of standard deviations of each dimension for the British pulation is not, in principle, easy. However, since the estimate needs to be curate only to 0·1 in. (and since the available material may itself be less ecise than it seems) the estimate for British males is derived from a sample ndard deviation by multiplying by

$$\frac{\text{British standard deviation of stature (Clements and Pickett, 1957)}}{\text{sample} \quad ,, \quad ,, \quad ,, \quad ,, \quad \text{(source 3 or source 5)}}$$

This procedure is not possible for the dimension buttock–back of calf, which is given in source 7, but not in source 3 or source 5. Since, however, source 7 agrees well with source 5, it seems reasonable to guess at its standard deviation of stature.

The estimates based on this reasoning are given in Table 4.

Table 6. Coefficients of variation: females

Source	4	11	12
Stature	3·8	3·9	4·0
Shoulder – seat	5·1		
Shoulder – elbow	5·6	5·7	

(*d*) *Estimates of standard deviations for the female population*

The available material is very defective. One group in source 1 had very few members and the other was highly selected for stature. The only reliable evidence is summed up in Table 6. The standard deviation for stature is the same for British females (source 12) as it is for British males (Clements and Pickett 1957). The only estimate of a British standard deviation that can be derived from the available material, shoulder–seat height (from source 4), is also identical with the value for British males. In view of this, it is tentatively assumed that each standard deviation for females is identical with the corresponding one for males (see Table 4).

§ 4. APPLICATION OF THE ESTIMATES TO SEAT DESIGN

The application of anthropometric data to design should be indirect as a rule; that is, the data should be used to specify the body sizes of people who test the equipment, rather than to obviate testing altogether. Natural postures are usually very different from the rigid ones most convenient for anthropometric surveys. Thus, there are several dimensions of chairs that are not directly governed by the dimensions specified in this paper. For instance, the minimum acceptable height and depth of the seat, the shape of arm-rests, and the angular measurements are examples of these. However, there are several other dimensions of chairs that depend directly on measurements quoted in this paper, for instance the maximum acceptable height and depth of the seat, and its minimum acceptable width; but it must be noted that buttock width is less than minimum thigh breadth. The measurements chosen for these dimensions will depend on what part of the population the chair is meant for.

Any application of the estimates in Table 4 must take into account that they refer to naked people. It is difficult to estimate the effects of clothing and shoes on measurements, particularly for women, but values are given by Roberts (1945) for American servicemen's clothing. The increases in measurements are:

Shoulder width	—	in indoor clothes	0·4 in.	with overcoat
Elbow width	0·6 in.	in indoor clothes	1·8 in.	with overcoat
Buttock width	0·6 in.	in indoor clothes	1·1 in.	with overcoat

Shoes cause an increase in the measurement of underside of thigh–floor. For men's shoes the mean increase is 1·1 in., and for women's shoes there are at present (1960) two modes, one at about 0·9 in. and one at 2·0 in. Slippers, both men's and women's, cause an increase of about 0·7 in.

The work on which this paper is based is part of the research programme of the Furniture Development Council, and is reported here by permission of the Head of the Research Department. It has been fully described in FDC Research Report No. 8, Anthropometric Data for Chair Designers, by H. D. Darcus, M. J. Merrick and D. M. Barkla. Thanks are due to the Dunlop Rubber Company for financial support for these investigations, and to the British Boot, Shoe and Allied Trades Research Association for information about shoes.

Il est assez difficile d'obtenir les rapports sur la plupart des études de mensurations humaines et les renseignements relatifs aux dimensions utiles pour la création de sièges sont généralement disséminés parmis d'autres sujets.

Le présent article rassemble les principales informations publiées concernant ces dimensions.

La plupart des populations étudiées ne sont pas, elles mêmes, importantes pour les créateurs britanniques mais une estimation des mensurations des jeunes adultes britanniques est déduite à partir des données publiées.

Berichte über die meisten Studien, die sich mit menschlichen Körpermassen befassen, sind nicht leicht zugänglich und die darin enthaltenen Angaben über die Masse, soweit sie sich auf Sitzmöbelentwurf beziehen, sind zumeist mit anderem Material vermischt. Diese Arbeit stellt die wichtigsten der veröffentlichten Massübersichten zusammen.

Die meisten der behandelnden Bevölkerungen sind als solche für britische Entwürfe nicht von Interesse, aber Körpermasschätzungen von jungen britischen Erwachsenen werden dem veröffentlichten Material entnommen.

REFERENCES

Åkerblom, B., 1948, *Standing and Sitting Posture* (Stockholm : A. B. Nordiska Bokhandeln).

Clements, E. M. B., and Pickett, K. G., 1957, Stature and weight of men from England and Wales in 1941. *Bri . J. prev. soc. Med.*, **11**, 51–60.

Darcus, H. D., and Weddell, A. G. M., 1947, Some anatomical and physiological principles concerned in the design of seats for naval war-weapons. *Brit. med. Bull.*, **5**, 31–37.

Fry, E. I., and Churchill, E., 1956, Bodily dimensions of the older pilot. *W.A.D.C. Technical* Report 56–459. Wright-Patterson Air Force Base, Ohio.

Geoghegan, B., Roberts, D. F., and Wilson, J. O. C., 1957, Miscellaneous functional measurements of naval personnel. *Royal Naval Personnel Research Committee* Report RNP 57/895. This document also includes Roberts (1957).

Hertzberg, H. T. E., Daniels, G. S., and Churchill, E., 1954, Anthropometry of flying personnel—1950. *W.A.D.C. Technical* Report 52–321. Wright-Patterson Air Force Base, Ohio.

Hill, A. Bradford, and Morant, G. M., 1945, Report on body measurements of 2200 tank personnel. *Military Personnel Research Committee* Report BPC 45/406. M.R.C., London.

Hooton, E. A., 1945, *A Survey in Seating* (Gardner, Mass. : Heywood-Wakefield Co.).

Kemsley, W. F. F., 1957, *Women's Measurements and Sizes* (London : H.M.S.O.) (Published by the Board of Trade, for Joint Clothing Council Ltd.).

McFarland, R. A., Damon, A., Stoudt, H. W., Moseley, A. L., Dunlap, J. W., and Hall, W. A., 1953, *Human Body Size and Capabilities in the Design and Operation of Vehicular Equipment* (Boston, Mass. : Harvard School of Public Health).

Morant, G. M., 1943, Preliminary report on body measurements of 2400 candidates for air crew. *Flying Personnel Research Committee* Report 538. Royal Aircraft Establishment, Farnborough, Hants ; 1947, Dimensional requirements for seats in R.A.F. aircraft. *Flying Personnel Research Committee* Report 682. R.A.F. Institute of Aviation Medicine, Farnborough, Hants.

Morant, G. M., and Gilson, J. C., 1945, A report on a survey of body and clothing measurements of R.A.F. personnel. *Flying Personnel Research Committee* Report 633 (a). Royal Aircraft Establishment, Farnborough, Hants.

Newman, R. W., and White, R. M., 1951, Reference anthropometry of Army men. *Environmental Protection Section* Report 180. Quartermaster Climatic Research Laboratory, Lawrence, Mass.

O'Brien, R., and Shelton, W. C., 1941, *Women's Measurements for Garment and Pattern Construction.* Misc. Publ. 454. U.S. Dept. of Agriculture, Washington.

Randall, F. E., 1947, Survey of body size of Army personnel, male and female : Methodology and general considerations (female). *Environmental Protection Section* Report 123. Quartermaster Climatic Research Laboratory, Lawrence, Mass.

Randall, F. E., and Baer, M. J., 1951, Survey of body size of Army personnel, male and female : Methodology. *Environmental Protection Branch* Report 122 (revised). Quartermaster Climatic Research Laboratory, Lawrence, Mass.

Randall, F. E., Damon, A., Benton, R. S., and Patt, D. I., 1946, Human body size in military aircraft and personal equipment. *Army Air Forces Technical* Report 5501, TSEAA–695–43. Air Materiel Command, Dayton, Ohio.

Randall, F. E., and Munro, E. H., 1949, Reference anthropometry of Army women. *Environmental Protection Section* Report 149. Quartermaster Climatic Research Laboratory, Lawrence, Mass.

Roberts, D. F., 1957, Descriptive anatomy of naval personnel. See Geoghegan *et al.* (1957).

Roberts, L. B., 1945, Size increase of men wearing various clothing combinations. Project No. 9. U.S. Army Medical Research Laboratory, Fort Knox, Kentucky.

NECK-MUSCLE TENSION AND THE POSTURAL IMAGE*

By F. P. Jones, F. E. Gray, J. A. Hanson and J. D. Shoop
Institute for Psychological Research, Tufts University,
Medford, Massachusetts

The ‘ postural images ’ of comfort, of correctness, and of height were studied in seven male subjects aged 16 to 21. The image was recorded photographically and electromyographically as the immediate response to a specific verbal stimulus. The response was quantified by measuring the angular relation of head to trunk and the change in electric potential of the sternocleidomastoid and upper trapezius muscles. Both sitting and standing postures were recorded. In addition, the effect on posture of effort (lifting), anticipation of movement, forced respiration, and standing on tip-toe was studied. The data were treated statistically.

§ 1. Introduction

In a previous paper Jones and O'Connell (1958) described a method for expressing dynamic posture in terms of the changing relation between head and trunk during movement. If this relation is used as the basic parameter, posture can be recorded as a patterned response to a specific verbal stimulus. The method provides quantifiable data—linear and angular measures—which can be treated statistically. We have used the method in combination with the electromyography of neck muscles to study a person's image of his ‘ most comfortable ’ posture, his ‘ best ’ posture, and his ‘ greatest height ’.

The idea of a ‘ best ’ posture seems to be held almost universally in our culture. Everyone we have asked claimed that he knew what good posture was, though he frequently apologized because he himself didn't have it. In defining best posture, individuals disagree as to what is and what is not an example of it. When asked, however, to sit or stand in his best posture, everyone whom we have observed has made some kind of response, modelled apparently after an image, consciously or unconsciously held. This image is usually contrasted with ‘ most comfortable ’ on the one hand, and with ‘ greatest height ’ on the other. Occasionally, a subject insists that the three mages are indistinguishable—that when he is in his best posture he is at his greatest height, and that this is also his most comfortable posture. To most subjects, however, at least two of the three are distinguished from each other.

The aim of the present study is to define these three images more objectively by measuring some of the changes that take place during the postural response. Of the many variables that might be used as measures, we selected the relation of head to trunk. This relation is central in the organization of posture in animals (Fulton 1949, Chapter 8) and we believe that it plays a rôle of similar mportance in human posture. We obtained simultaneous records of the elative positions of the head and trunk and of the electric potential in two muscles of the neck, the upper trapezius and the sternomastoid.

The sternomastoids are the two oblique muscles on either side of the neck. The two parts of the upper trapezius form a triangle running from an apex at he occiput to a base line drawn through the seventh cervical vertebra.

---

* This study was supported by a PHS research grant from the Division of General Medical ciences of the United States Public Health Service.

According to Gray (1948) the two sternomastoids " acting together flex the vertebral column, bringing the head forward " (p. 371), while the two sides of the upper trapezius " draw the head directly backward " (p. 423).

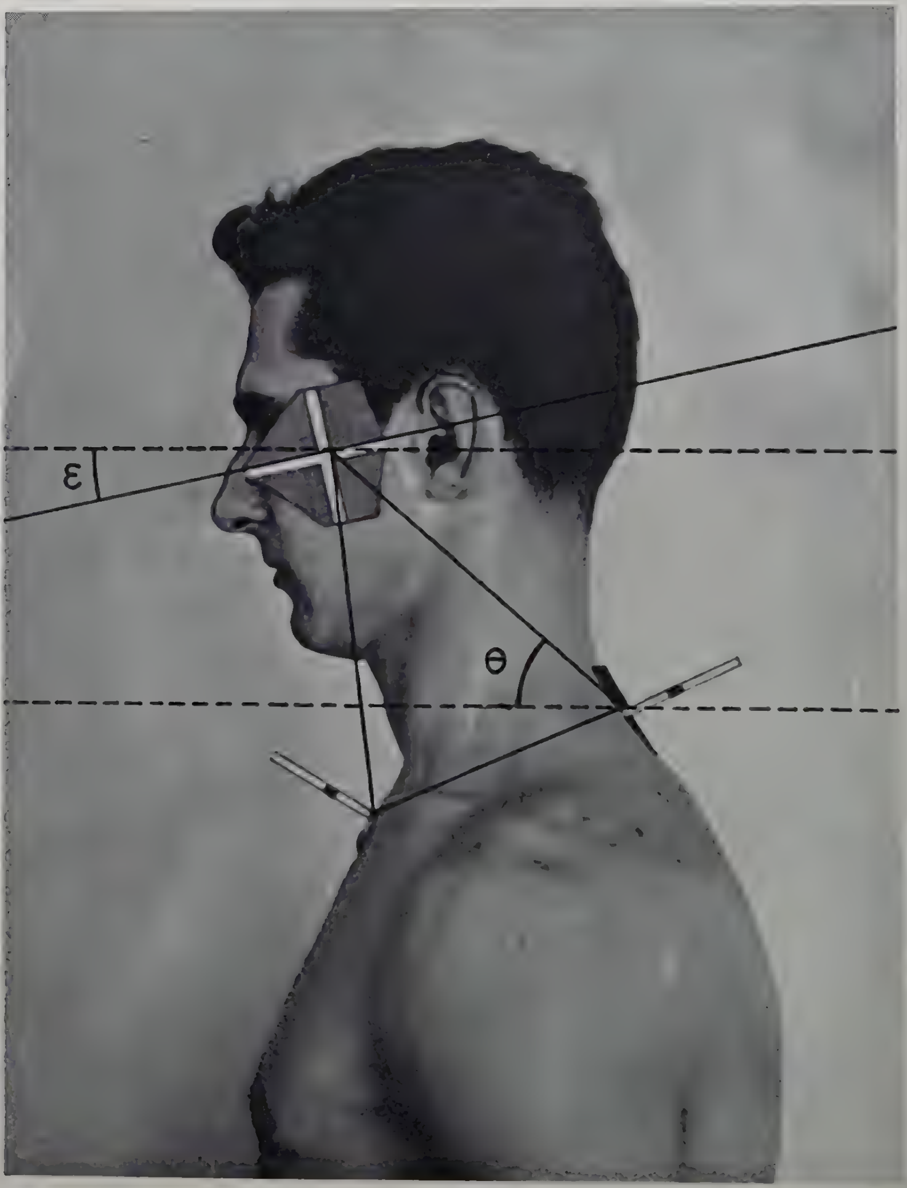


Fig. 1. Subject with markers in place and triangle diagram constructed. Broken lines indicate the horizontal.

Woodworth and Schlosberg (1954, p. 174) called attention both to the key rôle played by the neck muscles in postural activity and to their importance as indicators of the general level of muscular tonus. In this study we confined our attention to postural activity alone, recording from the two neck muscles and from one back muscle, the erector spinae.

§ 2. METHOD

The subjects were seven male students ranging in age from 16 to 21. All of them were employed for the summer by Tufts University, six of them in maintenance work.

2.1. *General Procedure*

The subject was seated on a flat stool in an electrically shielded room. The relation of the head to the trunk was marked by two light plastic rods attached over the seventh cervical vertebra and the sternal notch, and by a small cross centred in the Frankfort Plane, midway between the tragion of the ear and the lowest point of the orbit (Fig. 1). Pairs of Grass electrodes two inches apart were attached to the skin with Duco cement. One pair was placed over the right sternomastoid half-way between the insertion and the sternal head. A second pair was placed over the trapezius half-an-inch to the right of the central ligament and half-way between the occiput and the seventh cervical vertebra. A third pair was placed over the right erector spinae two nches below the lowest point of the trapezius. In addition, there was an earth electrode over the right scapula.

Muscle potentials were amplified by an Offner Type A Electroencephalograph placed outside the shielded room. Two integrators were inserted between the preamplifiers and the power amplifiers. Their output was fed nto two channels of the Offner. Direct muscle activity and integrator output were recorded by an inkwriter on paper moving at 2·5 cm/sec. The integrator scores appeared on the record as a series of condenser discharge ' pips ' proportional in rate to the amplitude of the signal. Photographs of the subject were taken by a Robot Star camera with a self-advancing film. The shutter was controlled by a solenoid operated by the experimenter from outside the booth.

The various procedures were explained to the subject in advance. He was asked, however, not to rehearse the movement but to wait for the signal and then make whatever seemed to him the appropriate response. The picture was taken as soon as the response was completed. The inkwriter was started before the signal and allowed to continue for approximately ten seconds after the picture had been taken. An auxiliary pen recorded the opening of the shutter and the signals given by the experimenter.

For six of the subjects, the series of procedures was repeated in order to obtain an integrated record of the third muscle. Consequently, 13 pictures were taken for each procedure. For statistical purposes, data from Subject 7 were tabulated twice, giving 14 observations ($n = 14$) for the postural patterns and for the sternomastoid muscle. Since only one integrated record was obtained from the upper trapezius, for this muscle $n = 7$.

2.2. *Specific Procedures*

2.2.1. *Sitting*

(a) *Postural images.* The subject, with left profile to the camera, was eated in whatever posture he chose to assume. The instructions were then iven in the following order : " Sit in your best posture " ; " Sit in your most omfortable posture ' ; " Sit at your greatest sitting height ".

(b) *Weight lifting.* To test the effect of effort on the three postures, the following procedures were used. The subject, seated as before, let his right hand hang by his side. A standard 2 kg weight was placed beside him. The weight was attached to an iron rod and swung freely from a wooden handle which the subject was able to grasp easily without changing his position. From each of the three postures, the subject on signal lifted the weight by bringing his arm forward to the horizontal. The rate of lifting was regulated by a metronome. The three procedures were then repeated with a weight of 4 kg. For each procedure a picture was taken when the arm reached the horizontal.

(c) *Ready to stand.* Four procedures were used to test the effect of anticipatory ‘ set ’. 1. The subject was instructed to move on signal from sitting to standing. In preparation for the movement he was asked to sit in his most comfortable posture. A photograph was taken just before the signal to get up. Both the preliminary posture and the movement were recorded electromyographically. 2. The procedure was the same as in 1, except that the subject was asked to get up more slowly than he was accustomed. 3. The same as 1, except that no instructions were given for the preliminary posture and the subject was asked to get up as quickly as possible after hearing the signal. In the protocols this was referred to as the ‘ reaction-time procedure ’. 4. The same as 1, except that in preparation for the movement the subject was asked to sit at his greatest sitting height.

2.2.2. *Standing*

The three postural images of ‘ best ’, ‘ most comfortable ’, and ‘ greatest height ’ were recorded for the standing posture in the same sequence as for sitting. In addition, the subject was asked (a) to take a deep breath and hold it ; (b) to breathe out deeply and hold it ; (c) to stand on tip-toes.

2.2.3. *Subjective reports*

At the end of the session, subjective reports were obtained by means of a short questionnaire in which the subject was asked for his definition of ‘ best posture ’, whether he distinguished it from the other two, and whether he had had any special training in posture.

§ 3. RESULTS

In answering the questionnaire, none of the subjects said he had had any special training in posture. Most of them were vague as to what they meant by ‘ best ’, and there was no agreement as to which of the three postures was most efficient. By conventional standards their postures ranged from excellent to poor. In view of this disparity, their responses during the experiment showed remarkable regularity.

3.1. *Postural Records*

To analyse the photographs, the negatives were projected and a triangle diagram constructed by connecting the two points on the trunk with each other and with the centre of the Frankfort Plane (Fig. 1). The plane itself was then drawn in to form a bar at the apex of the triangle. Coordinates for the three points were determined, and the angles made with the horizontal by the cross bar (ϵ) and one side of the triangle (θ) were measured. Means were btained for all of these measures. The final data were then replotted to give diagrammatic expression of the mean postures (Figs. 3 and 4). Means and andard deviations for the angles ϵ and θ (which measure the head's rotation nd its relation to the trunk) are given in Table 1. t-tests were made of the ifferences between the means in comparable procedures. Values of t and gnificance levels are given in Table 2.

Table 1. Means and Standard Deviations.

	Procedure		Angle (degrees)				Muscle (μv)			
			ϵ		θ		Sternomastoid		Upper Trapezius	
			Mean	SD	Mean	SD	Mean	SD	Mean	SD
1.	Sitting	MC	−13·64	16·96	35·72	14·99	7·16	2·38	15·05	10·77
2.	,,	B	− 1·42	10·57	50·48	9·17	10·34	6·55	9·44	7·55
3.	,,	GH	9·14	10·98	56·94	7·41	25·88	18·02	9·62	5·50
4.	Lifting 2 kg	MC	−12·84	17·00	40·12	15·77	26·23	6·05	70·40	29·29
5.	,,	B	− 1·67	10·33	51·83	9·18	31·32	11·09	51·24	20·91
6.	,,	GH	8·55	10·46	61·11	8·05	39·26	12·31	56·91	30·03
7.	Lifting 4 kg	MC	−10·64	21·59	41·25	19·82	46·23	9·88	122·97	34·88
8.	,,	B	4·19	9·80	58·35	8·33	45·06	9·46	95·85	38·99
9.	,,	GH	10·95	10·94	63·77	9·59	57·74	15·35	99·73	44·40
0.	Ready to stand	MC slow	−24·35	22·68	25·11	18·96	7·35	4·83	21·35	14·29
1.	,,	MC norm.	−23·44	24·41	24·56	21·04	10·15	5·22	21·55	14·17
2.	,,	RT	−23·03	20·29	23·78	16·64	7·04	2·38	22·98	11·57
3.	,,	GH norm.	7·08	9·73	55·61	5·93	20·51	18·48	6·27	4·26
4.	Standing	MC	−14·71	21·65	41·01	16·49	7·69	4·22	13·10	8·76
5.	,,	B	−3·16	16·51	51·06	13·07	15·39	11·19	12·98	11·69
6.	,,	GH	5·95	11·32	57·81	6·41	30·56	22·74	12·20	6·59
7.	,,	TT	− 2·56	7·82	47·29	11·55	16·03	8·87	7·81	3·86
8.	,,	Inh.	− 4·88	7·69	49·86	4·44	25·41	15·26	12·65	7·73
9.	,,	Exh.	−17·64	14·17	41·31	8·41	9·70	6·72	15·11	8·85

3.2. *EMG Records*

A continuous sample of the EMG record, taken while the subject moved om his most comfortable sitting posture to his greatest sitting height, is own in Fig. 2. In the top line, the first mark indicates the signal to the bject to move ; the second mark indicates the signal from the camera utter as the picture was taken. The second line carries the direct record om the right erector spinae. The third and fourth lines carry the raw uscle and the integrator count from the right upper trapezius, the sixth and venth from the right sternomastoid. Calibration records for 30 and 100 μv e shown on the right. It should be noted that the two integrators follow fferent scales.

Because the signal from the heart was so large in the lead from the erector inae (amounting to 30 μv or more) we did not use the integrator score from is muscle in our calculations. (It might be noted that the behaviour of the ector spinae did not distinguish the three postural images as did that of the o neck muscles.) In the neck leads, the heart played a much smaller role, nging, according to our best estimate, from $6\frac{1}{2}$ μv down to almost nothing.

Mean muscle activity was calculated by counting the number of pips in e integrator record over a 5 sec period and converting this number, by eans of calibrating signals, to the average equivalent in peak-to-peak μv/sec.

The final figure was obtained by subtracting the estimated ECG from the μv total.

The behaviour of the two neck muscles in the various procedures is shown diagrammatically in Figs. 3 and 4. Means and standard deviations for both muscles are given in Table 1 ; *t* scores and significance levels for differences are given in Table 2.

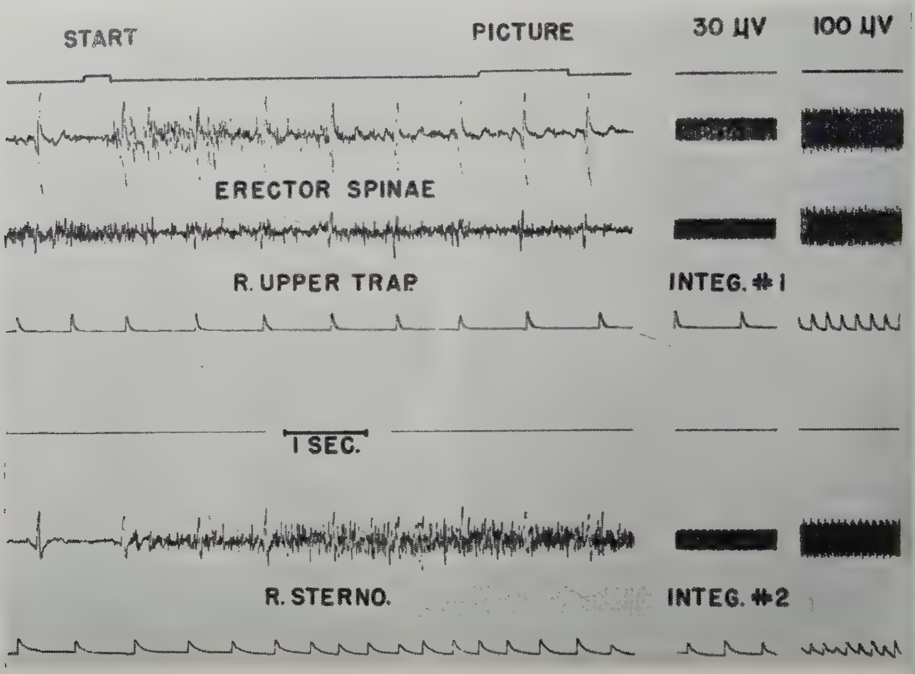


Fig. 2. Sample electromyogram. Movement from ‘ most comfortable sitting posture ’ to ‘ greatest sitting height ’.

Table 2. Differences between Means: *t* values and significance levels.

Procedure	Angle				Muscle			
	ϵ		θ		Sternomastoid		Upper Trapezius	
	t	sig.	*t*	sig.	*t*	sig.	*t*	sig.
1 vs 2	3·09	p 0·01	4·19	p 0·01	2·42	p 0·05	3·95	p 0·05
2 vs 3	4·61	p 0·001	3·92	p 0·01	3·77	p 0·01		
1 vs 3	6·84	p 0·001	6·55	p 0·001	4·03	p 0·01	1·72	
1 vs 10	2·39	p 0·05	2·95	p 0·05			2·66	p 0·0
1 vs 11	1·87		2 54	p 0·05	2·37	p 0·05	2·34	
1 vs 12	1·75		2·51	p 0·05			2·22	
4 vs 5	2·14		2·28	p 0·05	1·71		3·92	p 0·05
5 vs 6	5·01	p 0·001	4·61	p 0·001	2·84	p 0·05		
4 vs 6	4·38	p 0·001	4·77	p 0·001	3·73	p 0·01		
7 vs 8	2·62	p 0·05	3·12	p 0·01	0·32		1·53	
8 vs 9	4·42	p 0·001	3·54	p 0·01	3·60	p 0·01		
7 vs 9	4·55	p 0·001	4·66	p 0·001	3·03	p 0·05		
14 vs 15	1·52		1·65		3·65	p 0·01		
15 vs 16	3·50	p 0·01	2·72	p 0·05	3·76	p 0·01		
14 vs 16	3·30	p 0·01	3·49	p 0·01	4·15	p 0·01		
18 vs 19	3·33	p 0·01	4·03	p 0·01	4·68	p 0·001		

For ϵ, θ and the Sternomastoid, $n=14$; for the Upper Trapezius, $n=7$.

§ 4. Discussion

The triangle diagrams and the bar graphs provide succinct definitions for ach of the 19 postures recorded. A glance will show how relationships among he indices (four postural and two muscular) change from one response to nother.

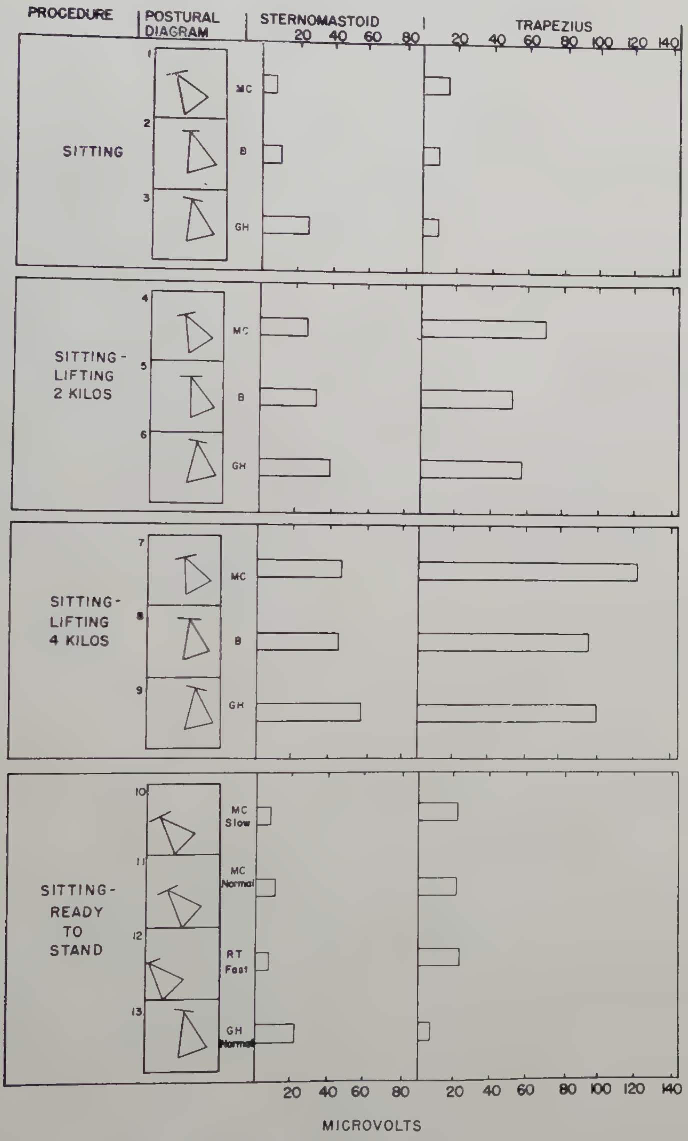


Fig. 3. Mean sitting postures (procedures 1 to 13).
MC=most comfortable ; B=best ; GH=greatest height ; RT=reaction time.

4.1. *The Nine Sitting Procedures*

The three postural images are sharply defined by the diagrams. In terms f the angles, ϵ and θ, each differs significantly from the other two at the 01 level or better. 'Greatest height' is the most stereotyped of the three responses. The chest is lifted and the head rotated backward at much the same angle in all of the subjects. ‘Best’, which falls between the two extremes, shows more individual differences. The greatest variation, however, is found in ‘most comfortable’, which ranges from almost completely erect ($\theta = 55°$) to almost completely collapsed ($\theta = 1{\cdot}5°$).

In the two lifting procedures, the three images are differentiated as before, with the greatest variation again in the most comfortable. Doubling the load makes little difference, except for a slight increase in ϵ which appears in each of the three postures.

A third measure which sharply differentiates the three postural responses is the behaviour of the sternomastoid muscle, whose potential increases as the subject moves from most comfortable to best and from best to greatest sitting height. The ability of this muscle to differentiate posture is retained in both of the lifting procedures, though its overall activity has increased greatly as a result of the lifting itself. An analysis of variance for the nine procedures showed that the postural differences were significant at the 0·01 level.

The postural activity of the upper trapezius followed a different (and often an opposite) course from that of the sternomastoid. It failed to distinguish the two erect postures from each other, but when the subjects moved from best to most comfortable, activity in the upper trapezius increased significantly. In this its behaviour is reciprocal to that of the sternomastoid. The reciprocal relation between the two muscles persists in the two lifting procedures, though in the 4 kg lift the difference in trapezius activity is not statistically significant. What the upper trapezius seems to be registering here is the postural change—the forward thrust of the head—that accompanies the subject’s attempt to relax in assuming a more comfortable posture. To test this possibility we asked one of the subjects to relax still further. When he complied there was a further increase in trapezius activity.

4.2. *The Four ‘Ready’ Procedures*

The effect of anticipation on the postural pattern can be seen most clearly in the reaction-time procedure (procedure 12). In getting ready to stand, the subjects regularly moved forward and down, increasing the forward thrust of the head. Though less marked, the pattern can still be seen in procedures 10 and 11, for which the subject was asked as a preliminary to sit in his most comfortable posture. Angle θ in both of these procedures differs significantly from the same angle in procedure 1, where there was no anticipation of movement. Only in procedure 13, where the subject was asked as a preliminary to sit at his greatest height, was this element of anticipation controlled. The pattern of 13 does not differ significantly from that of 3 and the two angles are virtually the same in both procedures.

It might be expected that anticipation of movement would produce a general increase in neck-muscle tension, but the expectation was not confirmed by our data. In four of our subjects, it is true, the sternomastoid showed considerable increase before the normal movement. But in the other three there was no increase at all, and before the slow movement and the reaction-time procedure (where an increase might well be expected) the mean actually decreased. The upper trapezius, which showed a slight but consistent increase, was probably reflecting the changes in posture only rather than nticipation of movement. It is significant that when the postural change as controlled there was no increase in either muscle.

4.3. *The Six Standing Postures*

The three postural patterns (most comfortable, best, and greatest height) re essentially the same for standing as for sitting. The most sensitive index the sternomastoid muscle, which distinguishes all three of the responses at ne 0·01 level. For the upper trapezius, on the other hand, the means were irtually the same for the three procedures.

Standing on tip-toes produced a pattern almost identical with that of best tanding posture. The response was more uniform, however, the standard eviation being smaller in all four indices.

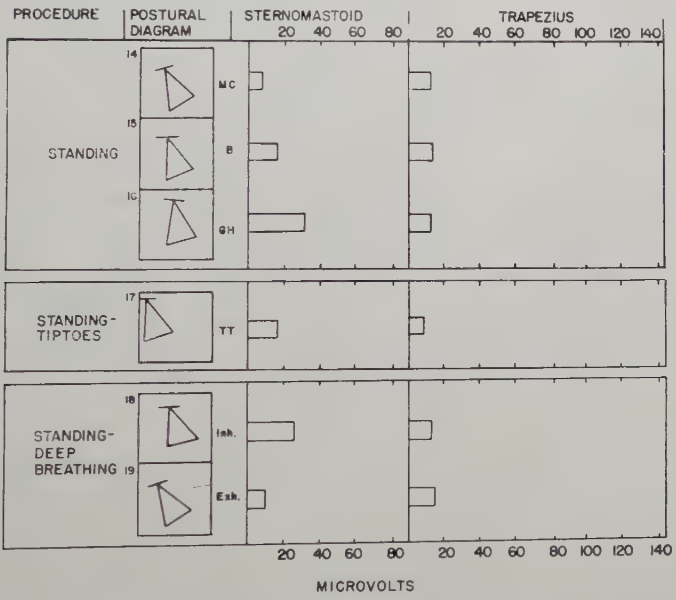


Fig. 4. Mean standing postures (procedures 14 to 19).
TT=tip-toes ; Exh.=exhale ; Inh.=inhale.

In the two breathing postures, the sternomastoid provides the best criterion ignificant at the 0·001 level). Potential in this muscle increases in inspiration, ecreases in expiration. The upper trapezius shows no significant change in ther procedure. The sternomastoid is one of the accessory muscles of spiration (Gray 1948, p. 391) ; it helps to lift the rib cage in forced inspiration. ortening to do so, it appears to tilt the head back as well, since there is an crease in ϵ that is significant at the 0·01 level. (Superficially, the patterns r inspiration and expiration resemble ' best ' and ' most comfortable ' spectively. The standard deviations for the two angles are much smaller the breathing postures, however.)

§ 5. Conclusion

In this study we have not attempted to define good and poor posture or set up any positive standards on which such definitions might be based. ur aim was to find out something about postural behaviour and estimate the extent to which this behaviour can be described in terms of stimulus-and-response. Though we were obliged to work with a small number of subjects, we believe that they were representative of their age and educational background. Within this group there was a wide range of postures, judged by conventional standards, but postural responses showed a high degree of regularity. The patterns, which seem to correspond to subjective images of comfort, correctness, and height, can conveniently be defined in terms of the relative position and angle of the head and the differential activity of two muscles in the neck.

On a étudié " l'image posturale " de ce qui constitue une position confortable, une position correcte, et une position de hauteur la plus grande. Les sujets étaient sept jeunes hommes entre seize ans et vingt-et-un ans. On a représenté cette image comme réponse immédiate à un stimulus spécifique et verbale. On a enregistré la réponse simultanément par la photographie et l'électromyographie ; et on l'a mesurée par les rapports angulaires entre la tête et le tronc et par les changements électriques dans le muscle sternocléidomastoïdien et dans la portion supérieure de trapèze. On a enregistré les réponses dans la position assise et dans la position debout. En outre, on a étudié les effets de l'effort du levage ; des mouvements anticipés ; de la respiration forcée ; et de la position de se tenir à la pointe des pieds.

On a traité les résultats de l'expérience avec les statistiques.

Man untersuchte bei sieben männlichen Personen im Alter von 16 bis 21 Jahren die innere Haltungsbilder für Bequemlichkeit, Richtigkeit und höchster Statur. Das Bild wurde jeweils als unvermittelte Reaktion auf ein spezifisches Reizwort photographisch und electromyographisch gleichzeitig aufgezeichnet. Die Reaktion liess sich quantitativ bewerten durch Messung des Winkels zwischen Kopf und Rumpf, und am electrischen Spannungswechsel im oberen Trapezius- und in den Sternocleidomastoideusmuskeln. Die Bilder wurden beim Sitzen sowie beim Stehen aufgezeichnet. Dazu untersuchte man die Einwirkung auf Körperhaltung bei Anstrengung (Heben), Sichrichten auf Bewegung, Tiefatmung und bei Stehen auf den Zehenspitzen. Die Daten wurden einer statistischen Behandlung unterzogen.

References

Fulton, J. F. (ed.), 1948, *A Textbook of Physiology* (Philadelphia and London : Saunders).

Gray, Henry, 1948, *Anatomy of the Human Body*, 25th edition (Edited by Charles Mayo Goss) (Philadelphia : Lea & Febiger).

Jones, F. P., and O'Connell, D. N., 1958, Posture as a function of time. *J. Psychol.*, **46**, 287–294.

Woodworth, R. S. and Schlosberg, H., 1954, *Experimental Psychology* (New York : Henry Holt & Company).

FATIGUE AND DRIVING

By A. Crawford

Road Research Laboratory, Department of Scientific and Industrial Research,
Harmondsworth, Middlesex, England

The problem of fatigue in driving is two-fold. It includes both the fatigue resulting from driving and the effects of fatigue, from whatever source, on driving. The study of fatigue in drivers is hindered by the fact that no reliable method of measuring fatigue or defining driving performance has yet been found. This paper reviews some of the research on operational fatigue, and discusses the source of fatigue in driving. It is concluded that stress is an important factor in the production of driving fatigue, and various methods of assessing the effects of the emotional arousal it produces are discussed.

§ 1. Introduction

Writers on the subject of fatigue have almost always found difficulty in arriving at a definition. The problem arises partly from the breadth and variety of the word's meaning in common usage and partly from the relatively small amount of direct research on the subject. Many of the effects of fatigue have been thought to account for the failure of experiments on other matters. Although some writers (Muscio 1921, Browne 1953) have questioned the acceptability of ' fatigue ' as a subject for study, it is agreed without exception that the effects are in themselves important and worthy of study. Floyd and Welford (1953) estimate that the cost of fatigue allowances in industry amounts to tens of millions of pounds per year. The contribution of fatigue to accidents is not known, but the nature of its effects leaves little doubt that the cost to the nation from this source must also be very high.

Many writers, including Bartlett (1953) and Broadbent (1953) define fatigue as deterioration in an activity as a direct result of being engaged in it. This definition is too narrow for the present purposes because it does not include the possibility of drivers being fatigued from other causes when they start to drive. That such a ' generalized ' fatigue is possible has been demonstrated in experiments which seek to find objective tests for fatigue such as an American study (Jones *et al.* 1947) of the effect of different periods of driving. In this experiment long-distance truck drivers were stopped and tested as they approached the terminus of a trip; the test results showed a general decline in several measures of psychomotor performance after long periods of driving. There are, therefore, two interrelated aspects of the practical problems of driving and fatigue, namely the fatigue arising from driving—operational fatigue—and the effect of fatigue, from whatever source, on driving. Whilst it may be possible to reduce the former, simply by cutting down the number of hours of driving, the effect of the latter requires study. For this, however, an objective measure of fatigue is necessary.

§ 2. Operational Fatigue

2.1. *Driving*

It is generally agreed that performance can be impaired by driving for too long a period but it has proved extremely difficult:—

(i) to define what is meant by driving performance,
(ii) to develop adequate techniques of measuring it,
(iii) to interpret the significance for fatigue of any signs of deterioration in performance,
(iv) to define the amount of deterioration which might reduce safety.

Measurement of discrete or isolated reactions throws little light on the deterioration of a complex skill. In order to examine changes in the patterns of a skill, techniques must be employed which, so far as driving is concerned, are at present only in the process of development. However, given these techniques, the most difficult problem still remains, that of administering the tests under conditions of tension and motivation comparable with those under which the skill is usually performed. Although research on ' fatigue in general ' has been exploratory and inconclusive, Welford (1953) can at least say of it that " the way has been cleared for a definite attack on the problem " and that we have reached the " end of the beginning ". Unfortunately this is not the case with driving fatigue where, since a basic knowledge both of fatigue and of driving performance is lacking, the beginning has not yet been reached. Nevertheless, a number of investigations have shown interesting results.

Reference has already been made to a study of the relationship between the number of hours of driving and certain psychological and physiological measures (Jones *et al.* 1947). This investigation covered 889 interstate truck drivers from three cities who were stopped just before reaching their destination, taken from their vehicles and tested. The results were compared with those for a control group, matched as to hours of sleep, who had done no driving. Of the measures used, hand steadiness, body sway, speed of tapping, and what the authors call ' reaction co-ordination time ' showed the greatest differences. There was a more or less consistent decrease in psychomotor performance with increase in the number of hours driving, but there was no sharp break in the results which might have indicated a particular moment of onset of fatigue. The value of the research was limited by the fact that the validity of the tests as predictors of poor driving or accidents was not known. However, the authors conclude that long hours of driving as well as the off-duty activities of the operators reduce ' efficiency ' and therefore the safety of driving. A similar but more intensive experiment reported by Ryan and Warner (1936) on six drivers tested after one hour and ten hours of driving also showed a demonstrable effect on steadiness, co-ordination reaction time and some other measures. The value of this experiment was also limited by the use of non-validated measures. Lauer and Suhr (1958), in attempting to validate laboratory experiments on driving by comparing their results with those of similar tests made during actual driving, showed that frequent rest pauses do much to reduce deterioration of performance due to prolonged driving.

In another type of investigation (McFarland and Moseley 1954), long hours of driving were imposed to bring about the fatigue condition, and the effect on driving was then noted. In this study the activities of two bus drivers on regularly scheduled passenger runs were directly observed and classified. Comparison of performance during the first and last half hours of the drive (the duration of the drive is not stated but may have been 3½ hours) showed onsiderable reduction in the number of steering wheel movements in the ast half hour (steering movements were made about 10 times as often as all other recorded movements together). A breakdown of the data into the number and kind of activities recorded for the first five minutes of each of even successive half hours showed no significant trends in any of the measures.

A further observational type of experiment, and one which appears to provide a very promising line of research, is the study of near accidents. An experiment of this type is that reported by Potts (1951) and by McFarland and Moseley (1954), in which seventeen long haul truck drivers on a total of 20 trips averaging 250 miles per trip, were accompanied by an observer. In all, 48 near accidents were reported by the observer, 46 per cent of them occurring within the first 2 hours and 67 per cent during the first 4 hours: in 72 per cent the trucks were travelling at 35 m.p.h. or faster. The numbers reported in each hour of the trip are shown in the following table:—

Hour of trip	1	2	3	4	5	6	7	8	9
Number of near accidents	11	11	6	4	4	4	4	3	1

These results, if they are valid and reliable, indicate that an increase in the driving period, far from being detrimental, leads to a reduction in near accidents. This startling finding is partly supported by a number of American studies on long haul truck accidents. McFarland and Moseley (1954) report that the accident prevention departments of two large insurance companies found that approximately 60 per cent of all long haul truck accidents in 1949 occurred during the first 3½ hours of driving; they also found that 'driver asleep' accident statistics indicated that the most dangerous time is 4 to 5 hours after the last sleep, and that the peak of the 'driver responsibility' accidents to interstate commercial drivers in 1938 occurred after 5 to 6 hours on duty. Unfortunately, all these data are open to the suspicion that they are reflecting the period of exposure of drivers to risk of accidents, and that before reliable assessments can be made it is necessary to know the distribution of the numbers of hours driven. These authors suggest that, as well as lack of rest, anxiety and other emotional states due to off-duty interpersonal relationships contribute to the pattern of accident behaviour. In other words, it appears that transferred fatigue has at least as important a part in accidents as operational fatigue.

Three 'hypnotic' effects of long monotonous journeys have been noted by American investigators (Haber *et al.* 1954). Most of the evidence for them is anecdotal but it is of sufficient volume to merit serious consideration. The effects are:—

(i) The driver's inability to appreciate his speed in terms of stopping distance when travelling at high speeds or when slowing to enter a restricted area.

(ii) A state of 'trance' brought about by traversing mile after mile of monotonous highway. For British drivers this is perhaps an experience to be expected in the future.

(iii) Hypnagogic hallucinations (McFarland and Moseley 1954), in which, after driving long distances, drivers imagine they see something on the road, and make emergency stops: the driver carries out the emergency stop, sometimes driving off the road, without recognizing that the situation is not real. The phenomenon occurs typically at night, while the vehicle is moving but while the driver's activity is at a low level. The content of an hallucination is most probably related to the driver's individual experience, but the frequency of occurrence seems to be very high in certain groups. All 33 of a group of long haul truck drivers interviewed reported having experienced these hallucinations, mainly at night (McFarland and Moseley 1954), whilst none of a group of 20 local truck drivers had done so. These drivers thought that the frequency of the hallucinations had been reduced considerably after the introduction of a shorter working day.

Unfortunately very little is known of the conditions producing these phenomena and no clear idea of their prevalence seems to have been formed by the investigators; their importance to the problems of long hours of driving in Britain is therefore unknown, but it is thought to be less than in the U.S.A., where distances driven are much greater.

2.2. *Complex Skills*

Since none of the experiments on driving show clearly the effect of fatigue as a result of driving, it may be useful to study the results of experiments on fatigue in relation to similar tasks. The most relevant work is that on operational fatigue and instrument flying, and Drew's conclusions (1940) are classical in this field.

Drew studied instrument-flight performance with a standard aircraft cockpit and simulated controls. Experimental sessions of 2 hours duration were divided into seven equal periods of intensive activity interspersed with three similar periods when, though complete relaxation was not possible, level flight was simulated; for some subjects the tests were extended to 6–7 h and their results confirmed those of the 2 h tests. Motivation and interest are said to have remained high throughout each test. There were six primary flying instruments, and when the subjects were fresh these were followed as an interconnected whole. Towards the end of the session, however, reactions to correct the individual misalignments shown on the instruments became more automatic and conventional, in the sense that the subjects used the kind of interpretation which first occurred to them regardless of correctness and became increasingly willing to accept lower standards of accuracy and performance. The instruments outside the immediate range of attention, a the edge of the display, were given less attention, and by the end of 2 h 60 pe cent of the subjects paid no attention to them at all. The timing scores improved initially, but, as the session progressed, began to deteriorate rapidly Finally, whenever the subjects were informed they were in the last stage of a run they showed an ‘ irresistible tendency to relax when home is in sight ’.

Bartlett (1943, 1951, 1952, 1953) has summarized his conclusions from hi own and succeeding experiments (e.g. Davis 1948) under the heading of ‘ Skil Fatigue ’. He distinguishes between fatigue resulting from hard physica work and fatigue from work calling for little muscular reaction but requirin persistent concentration and a high degree of skill. According to him, skille tasks, such as driving, require complex, co-ordinated and accurately time activities as well as simple repetitive movements. One essential characteristi f skill fatigue is that although the subject is physically capable of performing he desired task, he does not actually carry it out correctly unless he takes particular care.

With the onset of skill fatigue, the subjects' ' timing ' is affected (Bartlett 953); parts of the cycle of operations are occasionally slowed and other parts ushed to compensate. At a second stage subjects ignore some of the elements n the task. It might be said that the field of display is ' split up ' according o the proportion of signals ' seen ' or ' unseen ' in its various parts: outlying lements are likely to be either ignored, as Drew demonstrated, or stressed t the expense of the closely knit, central elements, and in consequence correct ctions may be performed but at the wrong time, while other actions are mitted altogether. At a third stage aches and pains are noticed; their ocalization is unimportant but, in common with many elements of the total ituation previously ignored, they reach consciousness and command response. Bartlett concludes that fatigue may be seen as a progressive widening of the ield of stimuli to which response is made, so that the subject's actions lose direction. Insight deteriorates fairly early and personal standards are lowered. Actions the subject would normally consider hazardous are undertaken with ittle concern when he is over fatigued (Davis 1948).

It is difficult to estimate how long it will take for each of these stages to be eached during driving. It seems unlikely that many drivers will advance ar in to thesecond stage without some incident occurring to stop them driving, or at least to make them alive to the dangers, and change their level of arousal.

Broadbent (1953) argues that a self-paced task is not likely to show the rrors of timing characteristic of the first stage of skill fatigue, but that if the ubject's working pace is dictated by the task he will be unable to compensate or momentary lapses and the task will be sensitive to fatigue. The ' clock est ' reported by Mackworth (1950) provides a task of the latter type: it equires the subject to watch a moving pointer and to report its very infrequent ouble jumps, with no possibility of compensating for a momentary lapse of ttention in which a signal is missed. Experiments with this test showed a ather rapid decrease in performance despite the small amount of overt action equired. This type of *vigilance* task seems to be one in which quick deterioraon is to be expected. In considering the way in which Barlett's concept of kill fatigue may be expected to apply to driving, it is at present a matter for onjecture how often the pace of driving is sufficiently forced for momentary pses to assume importance.

While laboratory experiments of this kind may indicate how the perceptual kills involved in driving are affected by fatigue, they cannot reproduce the ersonal contact between drivers, the competitiveness and the frustrations of he real-life situation. On the road a driver is trying to remove himself from he present situation and to reach his destination, probably as quickly as ossible. Such conditions lead to emotional responses which may dominate is behaviour and affect the rate of deterioration of his performance.

§ 3. THE SOURCE OF FATIGUE

There is a vast difference between the effect of repeated contractions of a uscle in response to stimuli from electrodes placed directly on it and the ndition of irritability, unusual sensitivity and irrational stubbornness, which is typically observed in those who have been engaged for a long period in sedentary work. Both conditions reduce efficiency; both are called fatigue. Between them lies a range of changes in the body chemistry and in overt behaviour the degrees of which seem to depend on the extent to which the central nervous system is involved in the task. That is to say, the more integrated the behaviour and, consequently, the more varied the possible sources of stress, the more gradually will the changes occur, the less will they depend on specifically physical effort, and the harder will they be to detect.

Except among truck drivers, who may man-handle their loads, gross musculature is not used to any substantial extent while driving, and drivers are not found to suffer from the symptoms of fatigue produced by severe exercise. On the other hand stresses may develop in heavy urban and holiday traffic, or in bad weather, or when road surfaces are difficult. In heavy traffic the driver's pace of driving is determined by the density and speed of the traffic stream and his willingness to conform. With the present road system the pace of driving is likely to be most firmly controlled by a traffic stream whose speed is usually fairly low, although sometimes rather higher: for example, along the first 20 miles of some of the main roads out of London the speed is high for the density of the traffic. It seems likely that there will be an increase in driving stresses from this source on the busier parts of the high speed roads of the future.

3.1. *Stress*

Emotional stress can have serious physiological results in subjects carrying out tasks in which there is little or no chance to correct errors; and stressful activities which do not lead to exhaustion—such as motor racing—produce severe depletion of energy reserves shown to be associated with deficiencies in hormone production of the adrenal cortex (McFarland and Moseley 1954). Studies of the galvanic skin response during driving have shown that there is considerable emotional response in traffic (Michaels 1960, Hulbert 1957), particularly when vehicles are turning, crossing, or merging, or are in close proximity to the test vehicle. A mild emotional response is not harmful provided that the driver is in reasonably good physical health, but repeated emotional arousal over a relatively short interval of time will raise a driver's general emotional level as one response builds on the arousal level of the previous one. As a result, an otherwise unexpectedly strong response may be evoked by a very slight irritation. Such behaviour is described as ' over-sensitive ' and the subject as ' fatigued '.

The second result of a sustained state of emotional arousal in a driver, over a short period, is that for some time afterwards it is more difficult than usual to elicit an emotional response. This may be regarded as another example of adaptation to stimulation; and in this case it will protect the driver from the severe depletion of energy reserves mentioned above. The effect may be described as a raised threshold of arousal. It will depend for its extent on the previous arousal level, and it will appear subjectively as depression, and loss of interest in the environment, and behaviourally as a reduced intensity or vigour of response to it.

One of the commonest sources of emotional arousal and stress is interpersonal relationships, and it is important to assess how much the stress from such sources contributes to the stress from traffic conditions. Whilst the stress of unresolved personal conflicts, fear, or anxiety is commonplace, the casual fleeting relationships with other drivers in the various traffic situations will also be a source of emotion to a varying degree, depending upon the driver's personality and temperament. Arousal from one set of causes can be added to that from another, as described above, and the one will always reduce success in overcoming the other. In extreme cases, when the strain is so strong as to interfere with sleep, a condition of chronic fatigue may develop. Such a condition follows prolonged periods of unpleasant emotion, and one of the best cures for it is physical exertion.

3.2. *The Sleep Mechanism, Emotion and Arousal*

In order to understand fatigue, then, it is necessary also to understand emotion and arousal. In view of this it will be no surprise that in recent assessments emotion, the arousal state, and the sleep mechanism are linked (Lindsley 1951, Magoun 1958). They are considered as different aspects of a single system having one continuum from deep coma via somnolence, states of relaxed wakefulness, bodily tension (aroused reflexly through lower levels of the spinal cord and the brain stem to unexpected and persistent external stimulation) to extreme excitement and emotional disturbance with loss of control and disruption of skilful activity.

Experiments have sometimes shown that there is very little deterioration in performance with loss of sleep. For instance Warren and Clark (1937) showed that after 65 h without sleep there was little reduction in performance level even though the subjects had to struggle to keep awake. One explanation of this is suggested by Laslett, who had no difficulty in securing volunteers for a similar experiment on lack of sleep, but had to prevail on his friends to act as subjects in the control series, with normal amounts of sleep (Seashore 1957); he concluded that the motivation of being in a difficult experiment helps to compensate for any deleterious effects of sleep loss.

Deterioration of performance resulting from lack of sleep is reflected in serial or continuous tasks like driving in traffic rather than in tasks which call for single actions and in which there are thus opportunities to rest. Subjects can arouse themselves momentarily, but if the test calls for continuous performance there is difficulty in maintaining arousal even for trials lasting a few minutes (McFarland and Moseley 1954, Seashore 1957). Many of the behavioural effects of forced wakefulness are of the same type as noted by Broadbent (1953) for vigilance tasks with rested subjects.

Malmo (1957) has shown that the experimental evidence supports the idea of an arousal or intensity dimension of behaviour as suggested by Lindsley 1951) and others, notably Duffy (1951), Hebb (1955) and Schlosberg (1954). The experiments, which were concerned with the relationship of physiological measures to the level of motivation (e.g. Stennett 1957, Survillo 1956), indicated hat the measures may serve to quantify this dimension of arousal.

§ 4. SOME POSSIBLE METHODS OF MEASURING 'STRESS' FATIGUE

In the introduction it was suggested that the problem of fatigue would be larified if an objective measure of fatigue could be found, and in the last ection it was suggested that a raised threshold of arousal following a period of stress is an alternative description for the kind of fatigue produced by driving. This section reviews the possibility of measuring the level of activation and the threshold of arousal.

In view of the fact that driving fatigue is identified with a raised threshold of arousal, the arousal value of any experiments or tests must be regarded as a parameter and either controlled or allowed for in the experimental design. For instance, if it is found that the subject's interest in the test itself is short-lived when he is fatigued but sustained for much longer periods when he is fresh, this will give a measure of the fatigue. A test has arousal value principally because it focusses the attention of the subject, but there are other contributory factors such as the personal satisfaction found in agreeing to co-operate, in the competitiveness, and in the self-knowledge gained from the results. For these reasons, until basic research on tests for fatigue is undertaken, experiments must be conducted as far as possible without the subject's knowledge.

4.1. *Physiological methods*

There can be no doubt that the most satisfactory index of fatigue would be a physiological measurement of some description. However, considerable practical difficulties in the recording techniques limit their usefulness at present.

The electroencephalogram has been shown by a number of writers (Lindsley 1951, Crawford 1955, Lindsley *et al.* 1949, Moruzzi and Magoun 1949) to present an activation pattern which may be found to give a measure of effort and of a subsequently raised threshold of arousal. Muscle spikes from electrodes placed on the pinna of the ear and from saline-moistened pads placed just above the eyebrows were found by Kennedy (1953) to vary with periods of alertness on a monotonous vigilance task. The change in electrical skin resistance (GSR) with sweating has been used for many years as a measure of emotional response. This method is regarded by Darrow (1936) as at least as good as any other measure, but it has the disadvantage that while *changes* are easy to record reliably, the *absolute* level of sweat gland activity is very hard to estimate. The list of current physiological measures of emotional arousal also includes blood pressure and volume, electrocardiogram and heart rate, respiration, skin temperature, pupilliary response, salivary secretion, pilomotor response, chemical sweating indices, analysis of blood, saliva and urine, metabolic rate, muscle tension, eye blink, and tremor. Of these, each is affected to a greater or lesser degree by sympathetic nervous activity but also by subsidiary reflex factors which cannot be fully controlled experimentally. At present, therefore, it is not possible to make use of the absolute values of physiological measures. Perhaps the most sensitive of them could be used with a standard arousing stimulus, and in this way the threshold of arousal gauged.

Schnore (1959) has examined the change in a number of these measures following stimulation under standard conditions and found not only that there were individual variations in arousal threshold but also that the measure showing the greatest change differed from subject to subject. However, Schnore makes the point that, despite the idiosyncratic differences in the patterns of results, any individual placed in an arousing situation will show an ncrease in most functions even though his general level may be different from he next person's. Comparisons made with these measures must therefore be f the amount of change in level of arousal in members of a single group of ubjects: the use of a control group is not applicable.

Apart from the disadvantage of the need for electrodes or their equivalent onnected to the subjects, and for sensitive amplifiers and recording equip- ent, it appears that physiological methods in general are more likely to be seful in indicating the changes in emotional stress during a task rather than he cumulative effect of it in the form of fatigue. However, there are signs hat the basal level of arousal may be identified, for example in comparing nxious with hysterical subjects by measuring the sedation threshold (Claridge 959, Shagass 1954). Practical application of the method to drivers might erhaps be approximated by giving them a sedative in a cup of tea after a eriod of driving and measuring the time before it takes effect.

4.2. *Psychological methods*

Mackworth's clock test, referred to in section 2.2., is an example of the ort of measure which may prove sensitive to a general level of fatigue. Other xamples are Kay's (1955) test of immediate serial recall, and the solving of airly difficult unfamiliar problems, such as those used by Welford *et al.* (1950). n Kay's test, subjects were required to press buttons in response to the ppearance of light signals, the correct button to be pressed being the one ssociated with, say, the third from most recent signal. This test proved ensitive to fatigue, and there were differential effects with age. In Welford's roblem-solving experiments the significant finding was that ' fatigued ' ubjects new to the task were less successful than those tested for the first ime when they were fresh. The first group when tested after recuperation till showed a poorer performance than the second group retested when they ere tired; the differential effect was thought not to be due to differences in he ability of the groups.

Perhaps the most sensitive tests are likely to be those which call, at least omentarily, for a high rate of arousal, and in which there is no opportunity or the subject to compensate for errors nor time for him to arouse himself if e is not ready. Various tests of this kind would be relevant to driving, but he tasks should be paced and should contain central and outlying elements, o that lowered standards of performance would show up. Preferably, the asks should also be easily administered without the driver having to leave the heel, but, presumably, with the vehicle halted. As a corollary to this type f method, fatigue might be measured by increasing a driver's work load by ngaging him in a secondary task while driving: an example of this technique hich required drivers to add digits or to listen to a string of numbers and otice when there was a repetition has been reported by Brown and Poulton 961) as a test of driving stress, although they did not attempt to use it to easure fatigue. When more than usual attention is required for driving, the rror score on the secondary task may be expected to increase; Brown and oulton found their task to be sensitive to traffic, and it is possible that a milar arrangement or an added secondary task would be sensitive to fatigue. owever, it might serve to reduce fatigue by contributing an added interest the task.

It is generally accepted that with the onset of fatigue the apparently less important aspects of a task will be sacrificed to keep the load of the task within the subject's capabilities. Thus among sedentary workers in their day-to-day existence the social courtesies are sacrificed first. In driving, these courtesies take the form of giving way to other drivers and to pedestrians, giving clear signals and, at night, dipping headlights. Reduction in these courtesies and increase of activities such as impatient horn blowing may be expected after long hours of driving. They may perhaps be used as measures of fatigue. Headlight dipping and horn blowing could be recorded automatically and on a large scale with simple apparatus and no-one in attendance. They have the advantage of not changing the task from driving to a special test, nor adding to it nor interfering in any way with comfort and normal driving habits. Such tests might, however, be difficult to arrange as it might not be possible to control the conditions and the measures would lose some sensitivity, and perhaps not show the effect at all, if the drivers had the added interest of knowing they were being tested.

§ 5. CONCLUSIONS

Fatigue resulting from prolonged activity at one task may be transferred to and measured by another. There are therefore two aspects to driving and fatigue. First it must be asked whether, and to what extent, driving leads to fatigue and secondly how much fatigue affects driving. It is entirely possible that driving creates a great deal of fatigue but, fortunately, is not very sensitive to it—or conversely that driving performance is very sensitive to fatigue but that fatigue from driving is limited. Research is required to distinguish between these possibilities and the results will indicate future research and administrative policy. At present a major difficulty in the way of such research is the lack of knowledge both about driving performance and about fatigue. Experimental techniques which do not yet exist or are only now being developed are required to study these basic problems.

It seems that the type of fatigue arising from driving is very different from that produced by physical exercise, and that it is caused largely by the stresses arising from traffic and other conditions, which produce varying states of emotional arousal. Repeated emotional arousal over a short period results in oversensitive behaviour showing strong responses to slight irritations, and if the period is prolonged, this is followed by a state of ' lowered vigour ' in which there is a reduction in intensity of response to the environment and a raised threshold of arousal. Either of these states is expected to increase the risk of accident both to the subject and to other drivers.

This paper is published by permission of the Director of Road Research.

Le problème de la fatigue du conducteur de voiture revêt un double aspect. D'une part, il y a la fatigue due au fait de conduire, d'autre part, il y a les effets de la fatigue—d'où qu'elle vienne—sur la conduite. L'étude de la fatigue chez les conducteurs est rendue difficile par l'absence actuelle d'une méthode fidèle d'évaluation du degré de fatigue et de critères de la performance lors de la conduite.

Cet article rapporte un certain nombre de recherches concernant la fatigue opérationnelle et fait porter la discussion su l'origine de la fatigue chez le conducteur. Il apparaît, en conclusion, que le stress est un facteur important dans la genèse de la fatigue du conducteur. On fait l'examen critique des diverses méthodes qui permettent d'évaluer les effets de la charge émotive du stress.

Das Problem der Ermüdung beim Kraftfahrer ist zweifach. Es umfasst sowohl die Ermüdung durch die Tätigkeit des Fahrens, als auch ermüdende Wirkungen irgend welcher anderen Herkunft auf das Fahren. Das Studium der Ermüdung bie Kraftfahrern ist dadurch erschwert, dass weder eine zuverlässige Methode der Ermüdungsmessung noch auch eine brauchbare Definition der Leistung des Kraftfahrers existiert. Die Vorliegende Arbeit sichtet einen Teil der Untersuchungen über Arbeits-Ermüdung und diskutiert die Ursachen der Ermüdung beim Kraftfahren. Es wird angenommen, dass Belastung (" stress ") eine wichtige Ursache der Fahr-Ermüdung ist. Verschiedene Methoden zur Feststellung der emotionellen Wirkung, die durch " stress " hervorgerufen wird, werden besprochen.

References

Bartlett, F. C., 1943, Fatigue following highly skilled work. *Proc. roy. Soc. B.* **131,** 247–254.

Bartlett, F. C., 1951, The bearing of experimental psychology upon skilled performance. *Brit. J. indust. Med.* **8,** 209–217.

Bartlett, F. C., 1952, Fatigue. *Lancet,* **262,** 868.

Bartlett, F. C., 1953, Psychological criteria of fatigue. In *Symposium on Fatigue* (Edited by W. F. Floyd and A. T. Welford) (London: H. K. Lewis), pp. 1–5.

Broadbent, D. E., 1953, Neglect of the surroundings in relation to fatigue decrements in output. In *Symposium on Fatigue* (Edited by W. F. Floyd and A. T. Welford) (London, H. K. Lewis), pp. 173–178.

Brown, I. D., and Poulton, E. C., 1961, Measuring the spare ' mental capacity ' of car drivers by a subsidiary task. *Ergonomics,* **4,** 35–40.

Browne, R. C., 1953, Fatigue, fact or fiction? In *Symposium on Fatigue* (Edited by W. F. Floyd and A. T. Welford) (London: H. K. Lewis), pp. 137–142.

Carmichael, L., and Dearborn, W. F., 1947, *Reading and visual fatigue* (Boston: Naughton and Mifflin).

Claridge, G. S., and Herrington, R.N. 1960, Sedation threshold, personality, and the theory of neurosis. *J. ment. Sci.,* 106, 445, 1568-1583.

Crawford, A., 1955, *Time of perception and the cerebral potentials.* Unpublished Ph.D. Thesis, in University of Edinburgh Library.

Darrow, C. W., 1936, The galvanic skin reflex (sweating) and blood pressure as preparatory and facilitative functions. *Psych. Bull.* **33,** 73–94.

Davis, D. R., 1948, Pilot error. *Air Ministry publication* A.P. 3139A (London: H.M.S.O.).

Drew, G. C., 1940, An experimental study of mental fatigue. Air Ministry. F.P.R.C. Report 227.

Duffy, E., 1951, The concept of energy mobilisations. *Psychol. Rev.* **58,** 30–40.

Floyd, W. F., and Welford, A. T., 1953, Editor's Preface to *Symposium on Fatigue* (London: H. K. Lewis).

Haber, H., Brenner, R., and Hulbert, S., 1954, The psychology of trip geography. *Bulletin* No. 91, Highway Research Board, N.A.S., Washington, D.C.

Hebb, D. O., 1955, Drives and the C.N.S. (Conceptual Nervous System). *Psychol. Rev.* **62,** 243–253.

Hulbert, S. F., 1957, Drivers' G.S.R.s in traffic. *Percept. Mot.* **7,** 4, pp. 305–315.

Jones, B. F., *et al.,* 1947, Fatigue and hours of service of interstate truck drivers. Quoted in R. A. McFarland and A. L. Mosely, *Human factors in highway transport safety,* 1954 (Boston, U.S.A.: Harvard School of Public Health).

Kay, H., 1955, Experimental studies of adult learning. Unpublished thesis in Cambridge University library, quoted in A. T. Welford, *Aging and human skill* (1958) (Oxford University Press).

Kennedy, J. L., 1953, Some practical problems of the alertness indicator. In *Symposium on Fatigue* (Edited by W. F. Floyd and A. T. Welford) (London: H. K. Lewis).

Kleitman, N., 1939. *Sleep and Wakefulness* (Chicago, Ill. University of Chicago Press).

Lauer, A. R., and Suhr, V. W., 1958. *Road adaptation of a laboratory technique for studying driving efficiency with and without a refreshment pause* (Chicago Ill.: National Safety Council).

Lindsley, D. B., 1951, Emotion. In *Handbook of experimental psychology* (Edited by S. S. Stevens) (London: Chapman and Hall).

Lindsley, D. B., Bowden, S., and Magoun, H. W., 1949, The effect upon the E.E.G. of acute injury to the brain stem activating system. *E.E.G. & Clin. Neurophysiol.* **1,** 475–486.

Lovatt Evans, C., 1956, *Principles of Human Physiology* (London: J. & A. Churchill).

McFarland, R. A., and Moseley, A. L., 1954, *Human factors in highway transport safety* (Boston: Harvard School of Public Health).

Mackworth, N. H., 1950, Researches in the measurement of human performance. *Med. Res. Coun. Lond. Spec. Rep.* No. 268 (H.M.S.O.).

Magoun, H. W., 1958, *The waking brain* (Springfield, Ill.: Charles C. Thomas).

Malmo, R. B., 1957, Anxiety and behaviour arousal. *Psychol. Rev.* **64,** 276–287.

MICHAELS, R. M., 1960, Tension responses of drivers generated in urban streets. Paper presented to 39th Annual Meeting, Highway Research Board, U.S.A. In press.

MORUZZI, G., and MAGOUN, H. W., 1949, Brain stem reticular formation and activation of the E.E.G. *E.E.G. & Clin. Neurophysiol.* **1,** 455–473.

MUSCIO, B., 1921, Is a fatigue test possible? *Brit. J. Psychol.* **12,** 31–46.

POTTS, C. R., 1951, A study of long haul truck operations. Quoted in R. A. McFarland and A. L. Mosley, *Human factors in highway transport safety*, 1954 (Boston: Havard School of Public Health).

RYAN, A. H., and WARNER, M., 1936, The effect of automobile driving on the reactions of the drivers. *Amer. J. Psychol.* **48,** 403–421.

RYAN, T. A., 1953, Muscular potentials as indicators of effort in visual tests. In *Symposium on Fatigue* (Edited by W. F. Floyd and A. T. Welford) (London: H. K. Lewis).

SCHLOSBERG, H., 1954, Three dimensions of emotion. *Psychol. Rev.* **61,** 81–88.

SCHNORE, M. M., 1959, Individual patterns of physiological activity as a function of task differences and degree of arousal. *J. exp. Psychol.* **58,** 117–128.

SEASHORE, R. H., 1957, Work and motor performance. In *Handbook of Experimental Psychology* (Edited by S. S. Stevens) (London: Chapman and Hall).

SHAGASS, C., 1954, The sedation threshold. A method for estimating tension in psychiatric patients. *E.E.G. & Clin. Neurophysiol.* **6,** 21.

STENNETT, R. G., 1957, The relationship of performance level to level of arousal. *J. exp. Psychol.* **54,** 54–61.

SURVILLO, W. W., 1956. Psychological factors in muscle action potentials. E.M.G. gradients. *J. exp. Psychol.* **52,** 263–272.

WARREN, N., and CLARK, B., 1937, Blocking in mental and motor tasks during a 65 hour vigil. *J. exp. Psychol.* **2,** 97–105.

WELFORD, A. T., 1953, The psychologist's problem in measuring fatigue. In *Symposium on Fatigue* (Edited by W. F. Floyd and A. T. Welford) (London: H. K. Lewis), pp. 183–191.

WELFORD, A. T., BROWN, R. A., and GABB, J. E., 1950, Two experiments on fatigue as affecting skilled performance in civilian aircrew. *Brit. J. Psychol.* **40,** 195–211.

WHITE, J. C., 1940, Autonomic discharge from stimulation of the hypothalamus in man. *Res. Publ. Ass. Nerv. Ment. Dis.* **20,** 854–863.

SOME EXPERIMENTAL STUDIES OF RADAR OPERATING

By D. Wallis and J. A. Samuel

Division of the Senior Psychologist,
Manpower Department, Admiralty

An account is given of experiments in which several critical features of the radar operator's task were examined. Experienced operators were used in each case, and special simulators were constructed to give realistic but controllable presentations.

Throughout the first study, subjects were required to undertake an auditory detection task at the same time as maintaining a visual search on the radar display. Each experimental watch lasted for 3 hours and was preceded and followed by a measurement of 'optimal' performance under psycho-physical test conditions. Results were analysed to compare the effects of 'continuous' and 'interrupted' radar operating, and to show how performance on watch is related to 'optimal' standards. Trends throughout a single session, and throughout five successive daily or weekly ones, were examined.

The second experiment was designed to assess the influence of target density (the number of contacts displayed simultaneously) and of blip-scan ratio (probability of 'paint') upon the detection of a new contact. Experimental sessions lasted for 30 minutes. An interaction between the two variables was established.

A third investigation is reported in which the eye-movement patterns employed in radar research were recorded by the electro-oculographic method. This was an exploratory study only, but it revealed consistent evidence of unsystematic visual scanning, with wide differences between patterns used by the operators.

The overall results and their implications for radar operating are discussed.

§ 1. Introduction

The primary function of a radar operator is to detect and identify signals which reveal the location of critical 'contacts'. A characteristic form of radar display is the plan-position-indicator (P.P.I.). It employs the circular face of a cathode-ray-tube to give map-like representation of the search area swept by a radar antenna. Signals appear as bright spots 'painted' on at the appropriate range and azimuth by a rotating radial line of light; they are likely to be masked by 'noise' or 'clutter' (e.g. returns from rain or land). The whole display brightens as the radial line revolves, but it gradually fades away until restored by the subsequent revolution.

Psycho-physical relationships between display variables and detection thresholds have been extensively studied under laboratory conditions. (Morgan 1952, has summarised the earlier work.) Necessary though such data are, however, they rarely permit more than the prediction of *optimal* performance, as defined by the operator who is fresh, alert, and fully aware of when and where a signal will appear on his display. Unfortunately, this gives at best an incomplete picture of what happens under real operating conditions. Since the pioneering studies of Anderson (1944) and Mackworth (1950), more attention has been devoted to the examination of *trends* in performance when operators are on watch for long periods and cannot anticipate the occurrence or location of radar contacts.

An account is given in the present paper of some experiments which sought to throw more light on the nature of radar operating, when it is practised under more realistic conditions than have usually been reported.

§ 2. PERFORMANCE DURING LONG WATCHES

THE first project was occasioned by a practical requirement to maintain a continuous watch on an airborne radar for several hours on end. At the same time a listening watch on a Direction-Finding (D.F.) system had to be maintained. Two operators, one officer and one rating, were available in the aircraft, but in separate cockpits. Each man could be supplied with a 5 in. P.P.I. display from the search radar, and with a headset to receive audible D.F. signals. Brief interruptions of his radar watch had to be permitted for one or other of the operators, to accommodate regular navigational computations. No break in the listening watch was permissible, however.

In this situation each man was faced with two simultaneous perceptual tasks. He scanned his P.P.I. for radar signals, and at the same time listened for characteristic audible signals. The two channels of information were entirely uncorrelated, and the signals in either one would be infrequent and noise-masked.

A full-scale mock-up of the aircraft cockpit was constructed and equipped with apparatus to simulate the radar and the D.F. listening channel. Radar echoes (2 mm diameter) were introduced against a controlled but changing background of noise and clutter; appropriate brightening and decay characteristics for the radial time-base and signals were shown, to accord with an antenna rotation rate of about 40 r.p.m. The auditory signals consisted each of 3 pure-tone pulses at 835 c/s in succession, each pulse lasting for 20 m/secs and separated from the others by $1\frac{1}{2}$ sec intervals. They were masked by a 'white noise' background which was fed continuously into the subject's headphones at an overall sound pressure of 75 dB re 0·0002 dyne / cm^2.

Within the cockpit, the ambient temperature and noise intensity were respectively held constant within ±3° of 62°F (dry-bulb) and ±2 dB of 66 dB re 0·0002 dyne / cm^2. Cockpit illumination was not measured in absolute units, but was maintained by means of external lighting at a constant subdued 'twilight' level.

2.1. *Experimental Procedure*

A high level of motivation was encouraged by means of an explanatory 'hand-out' and a personal interview of each subject before the experiments began. Radar and D.F. signals were individually demonstrated, and 5–10 min practice was given, with knowledge of results. Instructions emphasised the need for continual alertness on watch, and that a response button placed below the P.P.I. display was to be pressed immediately a fresh radar or auditory signal was detected.

The design of experiments was built around a 3-hour watchkeeping session in the cockpit, during which a pre-arranged programme of signals unfolded for each subject. There were always 10 separate radar signals and a further 10 auditory signals during each hour, dispersed irregularly and introduced without warning; signals never occurred simultaneously on both channels. Radar signals were repeatedly 'painted' on each rotation of the radial line until they were perceived. Auditory signals were introduced at an intensity below the threshold for detection in noise and increased by 1 dB steps every 4 secs until responded to. A confirmatory verbal report was required after the initial detection response: in the case of a radar echo, this report had to specify the spatial location of the displayed signal.

Subjects always attended in full flying kit, which included the standard aircrew headphones and microphone for the communication of reports and instructions to and from the experimenters. A short period of familiarisation with the procedure was allowed before the first experimental session, whilst every session was preceded by 5 min of formal practice.

Immediately after this practice, measurements were taken of the subject's ' optimal ' performance. For radar operating, this was defined as the average response time to 10 successive radar echoes presented with a prior alerting signal. Optimal listening performance was measured as the average signal-to-noise ratio at which 10 successive audible signals were just perceived, again with a warning before each signal occurred. Similar measurements of ' optimal ' detection were repeated immediately after each 3-hour session ended.

For the first experiment a group of 10 operators experienced two complete sessions in the cockpit. During one session both radar operating and the D.F. listening watch were entirely continuous. In the other session radar operation was interrupted for regular and specified periods totalling 15 min during each of the three hours. Exercises in navigational computation and plotting were taken up and continued in the cockpit throughout these breaks, whilst the subject tried to maintain his listening watch. Table 1 shows how the signals were deployed during these ' interrupted ' radar sessions.

Table 1. Plan of each Successive Hour During a 3 hour Session (' Interrupted ' radar Watch)

Serial Minute Number	Operating Tasks	Number of Signals
0–5	D.F. listening + Navigation	1 D.F.
6–20	,, + Radar Operating	2 D.F. + 4 Radar
21–30	,, + Navigation	2 D.F.
31–60	,, + Radar Operating	5 D.F. + 6 Radar

The second experiment continued with the same subjects, each of whom now underwent five sessions of ' interrupted ' radar. An interval of one week separated each session. Finally a third experiment was conducted with the remaining 12 subjects, who were given five sessions separated by one day only.

2.2. *Results*

Only the main findings will be reported here. Details of the breakdown of data and statistical analyses have been given by Wallis & Samuel (1956). So far as radar operating is concerned, the experiments allowed a comparison between continuous and interrupted working, and an assessment of trends in performance under the latter condition throughout successive daily or weekly spells of operation.

Taking the session as a whole, radar performance was significantly inferior when it was continued without a break. Moreover, the subjects were less efficient during the second hour than during the first; they recovered partially during the final hour, but only at the expense of a corresponding decline from the level of efficiency at which their simultaneous listening watch had been maintained. The only surprising feature of the downward trend in radar detection was that it did not appear till after an hour. This is much later than was expected on the basis of earlier findings from laboratory studies of vigilance (Mackworth 1950). However, the stress of continuous radar operation produced an adverse side-effect, since the average intensity at which the D.F. signals were just perceived was significantly higher throughout each successive hour, among the subjects who scanned their displays without a break.

During the second and third series of experiments, when the radar operating was always ' interrupted ', the advantages of this mode of operation were clearly confirmed. With one exception, various analyses of the results within and between sessions failed to reveal any trends, either in average level or variability of response. The only sign of accumulating stress was an increased variability of detection time among those operators who had completed three successive daily sessions. The removal of any tendency towards a progressive decline in operating standards as a watch proceeded, was anticipated on the basis of Mackworth's results. It confirmed the advisability of breaking up a radar watch into periods not exceeding one hour. Even so, it must be noted that occasionally there were brief but severe lapses of attention which led to abnormally long delays before a radar signal was detected. These lapses were just as liable to occur early in the session as later, as is evident from Figure 1.

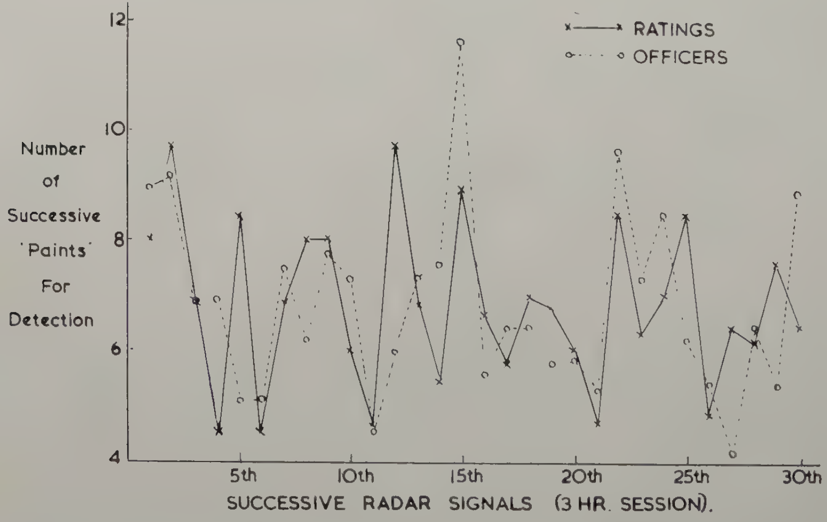


Figure 1. Detection of echoes: performance through ' interrupted ' watches of 3 hours duration. (Data from 12 subjects, averaged over 5 daily watches.)

It is interesting to note that the efficiency of the auditory D.F. watch paralleled that of the radar watch in all these later experiments. Not only were there no signs whatever of any changes in the average detection thresholds for auditory signals, but individual consistency was markedly high. Only rarely did any operator fail to detect a signal at an intensity within ± 2 dB of his own average response level. Reports from the subjects alleged that, subjectively, they were able to give all their attention to radar operating and could respond quite ' automatically ' to the sudden intrusion of a relevant auditory stimulus.

Immediately before and after each prolonged watch, measurements had been taken of how well the subject could detect when he was warned that a signal was imminent. These pre-watch and post-watch indices of ' optimal ' performance did not differ significantly. It was clear, therefore, that the intervening ordeal of three hours in the cockpit induced no ' fatigue ' of sensory functions. Yet among all subjects there was a marked fall-off in performance from this optimal level *as soon as* the proper session began; i.e., although the radar operators maintained a constant average level of performance throughout successive hours, they *never* managed to perform as well as they were demonstrably capable of doing when specifically alerted. The results are shown in Figure 2.

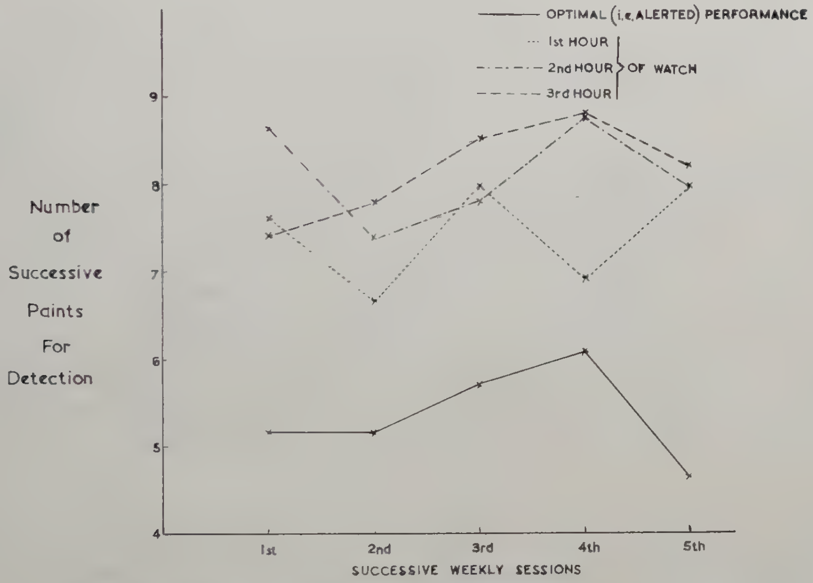


Figure 2. Comparison between optimal performance and detection during ' interrupted ' watches. (Data from 10 subjects).

This is an effect which has been noted in subsequent radar experiments, even with much shorter periods of operation. It occurs no matter how conscientiously the operator persists at his task, and seems to be related to the irregularity, infrequency and unpredictabliity with which signals actually appear. The same limitation, first demonstrated by E. Elliott at the Admiralty Research Laboratory (Elliott 1959) shows up in auditory watchkeeping tasks; as indeed it did in the synthetic D.F. listening watch employed in our experiments.

Another characteristic was revealed by analysing the responses to echoes displayed in various parts of the P.P.I. Several investigators have noted that some parts of the display tend to be ' favoured ', to the detriment of an efficient search when there is an equal probability of echoes appearing anywhere (Baker 1958). We found that detection was noticeably less reliable at long ranges, where a larger area for visual scanning is imposed by the P.P.I. form of display. This, as shown in Figure 3, was so even when the operator was fresh and alert, as in the pre-watch tests mentioned above ; and it was aggravated during normal search conditions. It suggests that the visual scanning patterns of radar operators may impart a bias against peripheral detections on a P.P.I., and that as time on watch increases, the bias is accentuated.

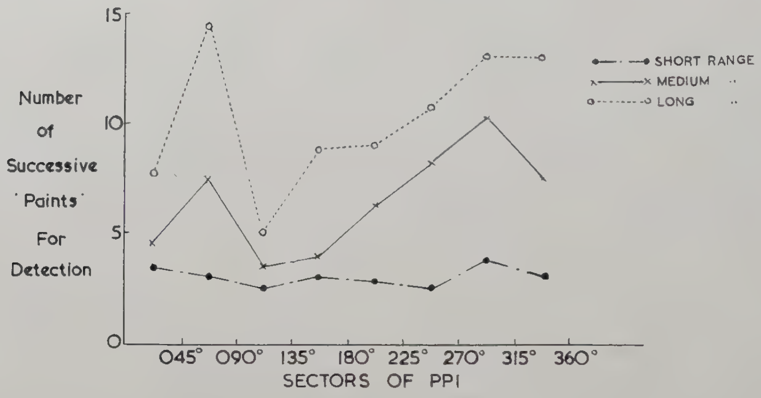


Figure 3. Detection in various display sectors and range bands. (Data from alerted subjects i.e. from pre-watch tests, averaged over 5 weekly sessions).

Finally, it should be emphasised that wide individual differences were demonstrated among operators. As groups, officers and ratings were not distinguishable; which is not surprising, since all the subjects seemed to be well-motivated during the experiments, and the operating tasks were easily comprehended. Yet although these subjects were all trained in radar and were well practised at the experimental task, they varied markedly and consistently from one to another within their respective groups. Significant positive correlations ($\tau = +\ 0{\cdot}72$, $p < 0{\cdot}05$ and $\tau = +\ 0{\cdot}70$, $p < 0{\cdot}01$) were obtained for each group of subjects between performance under ' optimal ' and under ' watchkeeping ' conditions. Taken together with other internal evidence, this suggests that the differences between operators during actual watchkeeping sessions may have been attributable to differences in ability to *search* the display effectively, rather than to differences in general *alertness*. (Even under the ' optimal ' conditions of the pre-watch measurements described previously, the subjects had still to *locate* the radar signals on their display.)

§ 3. Radar Reporting on Numerous Contacts

If the radar display is regarded as imposing a mental ' load ' on the operator, then clearly this load will vary within wide limits, according to operational circumstances. At one extreme there may be a very low incidence of ' relevant' signals, against a changing but perceptibly-distinct background of noise and other masking phenomena. Here the operator's load may seem to be minimal; yet the task is exacting, since it demands constant vigilance, systematic searching of the display, and the ability to discriminate an unpredictable target at the lowest possible signal-to-noise ratio. The investigation just described was the first of several into this ' search ' phase of radar operating.

At the other extreme the operator's load is qualitatively as well as quantitatively different. When there are many relevant targets already displayed, and others are expected, the maintenance of alertness presents little difficulty. Identifying a new target, however, becomes progressively harder for the operator; for he has to remember the whereabouts of 'old' targets and go on watching them, besides searching for new ones.

The following experiment was undertaken to estimate how great a load, in terms of target density (the number of targets painting simultaneously), lies within the capacity of a skilled operator. In practice, of course, target detection and identification are more difficult if probabilities of paint are low. For this reason, blip-scan ratio as well as target density was systematically varied. (The blip-scan ratio expresses the probability of a contact 'painting' on the display; empirically, it is found by dividing the number of successive antenna rotations into the number of occasions when a 'paint' occurred.)

A new similator was constructed to provide a controllable but realistic radar presentation, this time with a data of 6 r.p.m. Six experienced operators acted as subjects, using a 5in P.P.I. display. Each man carried out nine experimental watches of 30 min duration, the watches being taken on successive days. During the first 5 min the radar picture built up to one of the three conditions of target density, which prescribed 5, 10, or 15 radar echoes to be showing throughout the watch. A further 10 new targets appeared singly at random intervals during the remaining 25 min. To preserve the required target density one of the existing targets would fade just after a new one appeared. All targets tracked slowly across the display. Three different blip-scan ratios, of 0·5, 0·75, and 1·00, were employed, so that each subject experienced every combination of target density and blip-scan ratio.

The operating task was to press a response button as soon as a 'new' target appeared, confirming the detection by a verbal report of its range and bearing. All target echoes were similar in appearance; they were about $1\frac{1}{2}$ mm in diameter and just bright enough to be clearly visible when directly pointed out.

3.1. *Results*

It was known from several unpublished reports that the probability of detecting radar targets was related directly to the size of the blip-scan ratio and inversely to the target density. In practice, both variables must be considered together, and it was suspected that the interaction between them may lead to a critical breakdown in effective operating.

Results were analysed with the object of revealing such an effect if it were present, using the criteria of completely undetected targets and 'false' detections (i.e. already-existing contacts which were incorrectly reported as new ones), as well as response times for correct detection, to assess operating efficiency. Fig. 4 illustrates the relations between the two variables and the response times for targets correctly identified as new detections. There certainly appears to be a trend for increased target densities and lower blip-scan ratios to degrade performance, and an interaction between these variables shows up clearly. Nevertheless, only the interaction proved to be statistically significant; a sharp decline in performance occured only when the target density was more than 10 and the blip-scan ratio less than 0·75.

Changes in blip-scan ratio, considered independently, had no effect upon omissions or false detections, whereas both these measures increased significantly at the highest target density. No interaction between the variables appeared here. The overall conclusion was that the efficiency of radar operating can be seriously jeopardised, particularly at low blip-scan ratio, if the operator is forced to discriminate among more than 10 simultaneous contacts. This was confirmed in later experiments where fast-moving aircraft targets were presented on a P.P.I. display at a data rate of 15 r.p.m.

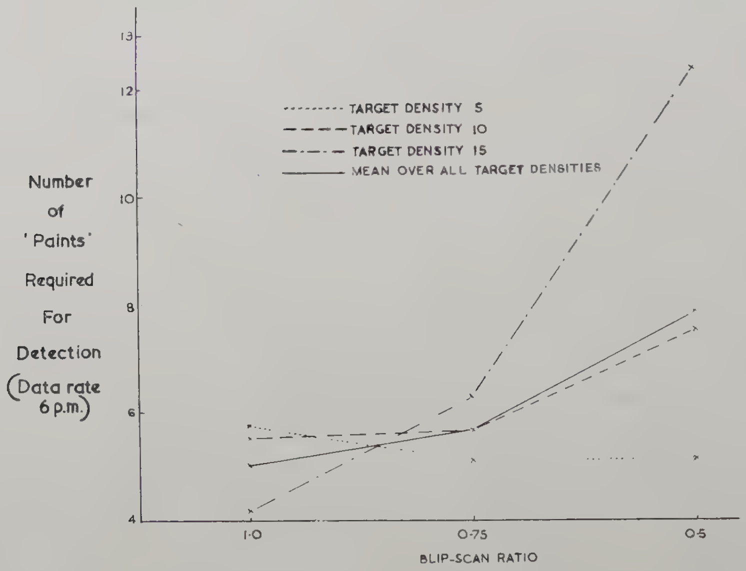


Figure 4. Effect of target density and blip-scan ratio upon detection of new targets.

A final effect which should be noted was the significant individual differences among these subjects, highly-skilled and experienced though they were, in the speed with which they could discern new targets. All subjects were prone to make false detections occasionally.

§ 4. Some Implications for Radar Operating

Before going on to mention work which is still in progress, let us sum up the main implications of these two investigations. Subsequent field trials using actual radar equipments under realistic conditions of operating, have produced results which support the following generalisations.

4.1. *Characteristics of Watchkeeping Performance*

It has become a classical doctrine that operators cannot remain alert during a surveillance watch, that they decline steadily, and that half-an-hour is as long as they should remain on duty without a break. Whilst there is some support for this view, the picture is really much more complex, and on the whole much less pessimistic.

In the first place, one must distinguish between ' low ' and ' high ' load conditions of reporting. Paradoxically, it is when the radar activity is minimal that efficiency is most likely to suffer. When the relevant signals are infrequent and cannot be precisely anticipated, which is characteristic of ' early warning ' and search phases of operation, circumstances are certainly unfavourable to human operators. They will tend always, even from the start of a brief watch, to respond less efficiently and consistently than might be predicted from a knowledge of how well they can detect signals which they can anticipate in time.

There is more to this than the simple assumption that operators ' doze off ' or are irresponsible. Unless alleviating measures are taken, the effect is present even though the personnel are well trained and sincerely striving to keep a vigilant watch. Fundamentally, it is explicable in terms of the psychological need which all human beings have for a ' lively ', perpetually-variable environment in order to conserve a high level of ' arousal ' (Bexton, Heron & Scott 1954) in the central nervous system. In the absence of variable stimulation from the physical surrounding, the sensory and nervous mechanisms are inevitably depressed. This is one reason why it is feasible to improve radar performance by introducing a less monotonous, more active, and perceptually more viable environment for radar operating than is often met with in practice. Experiments in Canada (Smith & Boyes 1957), confirmed in most respects by one of our colleagues (N. S. Kirk) with Service equipment, have shown that a moderately high level of illumination in radar compartments, together with careful setting up of displays, is beneficial to efficient reporting. Other measures, too, are practicable for creating a stimulating environment in which to operate.

An associated factor in ' low load ' conditions is the apparent regulation of performance by the frequency with which relevant contacts appear on the display. It is as though the operator's readiness to respond was conditioned by an estimate of the likelihood that anything will in fact be presented. This subjective ' expectancy ' is presumably based on experience from the onset of a watch, and will be least when the actual frequency of signals is minimal. Various investigators have shown that the efficiency of detection can be raised by artificially raising the signal rate, i.e. by periodically injecting ' false ' signals. (Deese & Ormond 1953).

An experimental system based on this idea was devised and tried out here (Wallis & Newton 1957), but in the form used its results were insufficiently promising to warrant its development. Nevertheless, in the light of knowledge gained in this study, it seems likely that a substantial aid to surveillance operations could be produced, in which artificial signals are added whenever ' real ' activity is excessively low. Precautions against the consequences of false alarms could certainly be built in. (It may be pointed out here that operators will produce a remarkably sensitive performance if they are encouraged to report anything which conceivably might be a contact, without waiting for the confirmation of successive ' paints '. This is tantamount, of course, to an artificial raising of the signal frequency, for it increases the probability of detecting real signals at the cost of an increase in ' false ' detections).

So far as ' high-load ' operating is concerned, we must recognize that although some continuous activity on the display is beneficial, there are upper limits to the amount of information with which an operator can deal. A number of variables influence these limits, so that one cannot generalize with any confidence. Circumstances which increase the irrelevant information, such as the presence of jamming and clutter, will naturally be detrimental.

So will those which reduce the relevant information, such as excessively low blip-scan ratios and signal-to-noise ratios. In any event the target density will be a limiting factor, and it is safe to assume that operators will be unlikely to cope effectively if there are as many as 10 contacts to be watched simultaneously on the display.

Finally, it is probable that ‘ fatigue ’ during a watch, in so far as this is manifested by a progressive decline in efficiency, is not a serious problem. It can be circumvented altogether by allowing a break in continuous operation after about half-an-hour, provided that the intervening period is occupied in a quite different kind of activity. Should it be operationally necessary, a watch which includes 10 min breaks at these intervals could be prolonged for three hours—though this is definitely not to be recommended as general practice ! We have confirmed several times that an hour's continuous watch can be maintained without a steady decline in performance when contacts are not less frequent than one every 10 min.

4.2. *Individual Differences in Skill*

Perhaps the most significant findings are those showing the marked variability which exists even among trained and experienced operators. This variability cannot be attributed to differences in visual acuity or general intelligence, at least among Service operators. It seems likely that personality factors may be influential, since radar operators certainly vary in the strength of their liking for the work and in the overt signs of stress or relaxation which they display on watch. There is a hint in Broadbent's (1958) account of vigilance research that the personality dimension of introversion-extraversion may be related to success in ‘ monitoring ’ occupations; but this has not so far been confirmed for radar operation.

Most of the research which is relevant to individual differences in radar operating has stemmed from the problems of training personnel in the perceptual skills involved. These training problems cannot be solved without a clear understanding of the radar operator's skill, one element of which seems to be the ability to carry out a systematic *search* of the visual display. The last part of the present paper will be concerned with a preliminary study of this factor.

§ 5. EYE-MOVEMENTS DURING RADAR SEARCH

Human beings tend to vary more among themselves at visual tasks than at auditory ones; and they vary still more when the visual task contains a large element of search, as in scanning a radar display. Earlier in this paper it was pointed out that the ‘ order of merit ’ produced by individual differences in the efficiency of prolonged radar watchkeeping, is substantially the same as that produced if the operators know *when* a target is about to appear although they do not know *where* to look for it on the display. These findings suggest that some of the variability between operators is directly attributable to varying degrees of skill in scanning the display.

The data illustrated in Fig 3 support the contention that some parts of a P.P.I. display tend to be ‘ favoured ’, even though the appearance of signals is equiprobable anywhere. Here is further evidence that patterns of visual scanning may incorporate undesirable biases away from a theoretically-optimal search, and may indeed be a cause of the surprising irregularities in the level of efficiency which appear from time to time during a watch. Unfortunately, explicit training in how to scan a display is not at present practicable, especially in relation to prolonged periods of systematic search, since no-one knows which methods are best or what operators are capable of.

Many operators seem to devise their own procedures. The majority claim that their usual method on P.P.I.s is to search just behind the radial time-base, but at data rates exceeding about 30 p.m. this becomes impossible, The subjects employed in our experiments made varied assertions about the patterns of search adopted, depending on the size of display and the data rate. The reliability of many of these claims may be doubted; but until recently there was no sufficiently-sensitive means of recording fixations and eye-movements, with which to examine the matter objectively.

However, several promising techniques have now been developed, one of which involves the use of equipment known as the Eye Movement Measuring Apparatus (Shackel *et al* 1958). This equipment, referred to as E.M.M.A., has been used in conjunction with the radar simulator devised for target density studies, to make preliminary observations of eye-movements during radar scanning (Samuel 1958). The method, known as electro-oculography, involves the pick-up and amplification of changes in potential difference between the front and rear of the eyeball, changes which occur as the eye-ball moves. With the E.M.M.A. equipment it is possible to observe fixations, and movements of the eyes as the point of visual fixation varies, in terms of the displacements of an oscilloscope spot. This display can be photographed at the same time as a monitoring display of the P.P.I. to which the experimental subject is attending.

A group of experienced operators have been ‘ monitored ’ during 30 min watches on a P.P.I. display. Their task was to search for new ‘ pips ’ appearing among those already present. A data rate of six per min was employed and the diameter of the display was 5 in. From the viewing distance of 17 in., the maximum angular deflection of the point of visual fixation was 17° across any diameter of the tube face. Certain restrictions were imposed on the operator by the use of electrode fittings near the eyes, but it was established that these did not detract from his ability to detect. (It is not verifiable, of course, that his eye-movements were entirely unaffected).

This investigation showed :

(i) There were wide differences between operators in the way they used their eyes to search the display, although a basically-circular pattern predominated.

(ii) Not only was there a general tendency to view the inner portion of the P.P.I. more closely than was appropriate—but a considerable proportion of the time was spent looking away from the display altogether.

Figure 5 gives a qualitative picture of what occurred in practice. It records the movements of one operator's eyes during six successive sweeps of the rotating time-base. Each ‘ frame ’ therefore accounts for a 10 sec period, during which the eyes moved between all the points connected. The order in which these movements occurred is not shown. The large number of movements certainly implies that the operator was trying hard to find a target, which in the present

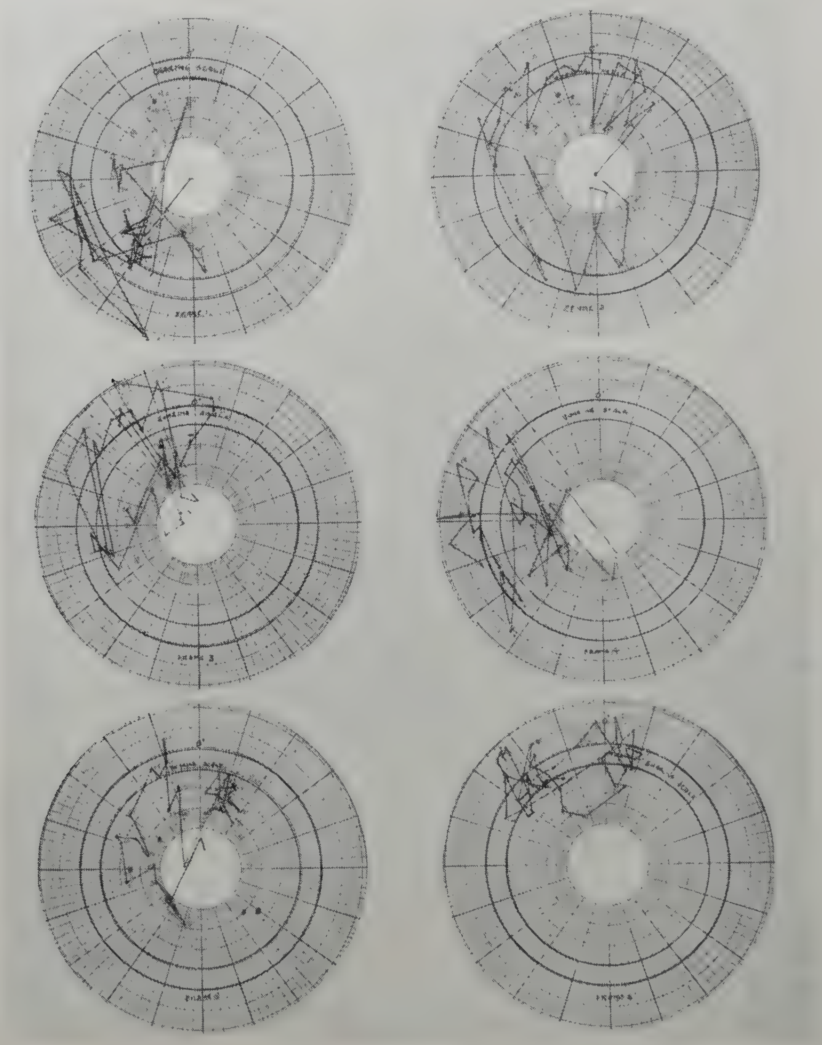


Figure. 5. Movements of the eyes during six successive rotations of radar scan (The larger black dots e.g. at 4 o'clock on frame 5, show where echoes painted).

case he seemed to expect somewhere in the upper half of the display. However, it is most unlikely that so intensive a search could be maintained for more than a few minutes. Nor is it likely that the more-or-less random nature of successive fixations afforded the most efficient or economic search pattern.

We cite this as only one example of the recorded patterns. Operators vary even with the same display; we suspect that display size, data rate, and time on watch, will all add to the variability. Rather than pursue the impossible path of recording what happens in all circumstances, it is necessary to go back to the beginning, as it were, to investigate just how effectively the eyes can detect marginal signals when these are displayed at varying angular deflections from the point of foveal regard (i.e. fixation point). This research is being conducted by our colleague N. S. Kirk. It is hoped to use the results as a basis for predicting what the most economic and efficient search patterns should be for different circumstances. Such information, if it should prove to be reliable in subsequent experimental studies, should be of great assistance not only to designers of future displays but also for the training of operators.

The two main experiments referred to here were carried out when the authors were attached to the Admiralty Research Laboratory, Teddington. Members of the Naval Motion Study Unit who took part included Messrs. S. J. Murch, W. D. Laurie and C. McCarthy (now with the Fleet Work Study Organization). Sincere thanks are due also to the Naval Officers and Ratings who acted as subjects in the various experiments.

The views expressed in this paper are those of the authors, and should not be taken as representing the official opinion of the Admiralty.,

Cet article donne un compte-rendu d'expériences dans lesquelles on a étudié plusieurs caractéristiques critiques du travail de l'opérateur radariste. Dans chaque expérience on a fait appel à des opérateurs entraînés et on a construit des simulateurs spéciaux permettant de réaliser des conditions de tâche réelle et contrôlable.

Dans la première expérience on a demandé aux sujets de réaliser une détection auditive tout en explorant par la vue l'écran radar. Chaque période de surveillance expérimentale durait 3 heures et fut précédée et suivie par une mesure de la performance " optimale " lors d'épreuves psycho-physiques. L'analyse des résultats porte sur la comparaison des effets du travail " continu " et du travail " interrompu " et montre comment la performance à une tâche de surveillance peut être mise en rapport avec des types " optima ". On a tenu compte du " trend " au cours d'une seule passation, de même qu'au cours des cinq jours ou semaines consécutifs.

La deuzième expérience devait permettre l'étude de l'effet de la densité des signaux (nombre de contacts présentés simultanément) et du rapport signal-balayage (probalité d'un " affichage ") sur la détection d'un nouveau contact. Une passation durait 30 min. On a pu montrer l'existence d'une interaction entre les deux variables.

Dans une troisième expérience on a étudié par une méthode électro-oculographique, le schéma moteur oculaire utilisé dans l'exploration de l'écran radar. Bien que cette étude ne fût qu'une étude d'approche, elle a permis de faire apparaître très nettement une exploration visuelle non systématique avec de grandes différences entre les schémas utilisés par les opérateurs.

La discussion porte sur les résultats généraux, ainsi que sur leur application à l'exploration par radar.

Es wird über Experimente berichtet, in denen einige kritische Besonderheiten der Aufgaben eines Radar-Beobachters untersucht wurden. In allen Fällen wurden erfahrene Beobachter herangezogen und besondere Simulatoren konstruiert, die wirklichkeitsgetreue aber Kontrollierbare Darbietungen geben.

Im ganzen ersten Versuch hatten die Versuchspersonen gleichzeitig mit der visuellen Suche auf dem Radarschirm akustische Signale zu entdecken. Jede experimentelle " Wache " dauerte 3 Stunden. Vorher und nachher wurde die " optimale " Leistungsfähigkeit unter psycho-physikalischen Testbedingungen gemessen. Die Ergebnisse wurden analysiert, um die Wirkungen " fortlaufender " und " unterbrochener " Radarbeobachtung zu vergleichen und um zu zeigen, wie die Leistungen bei der Beobachtung sich zu den optimalen Test-Standards verhalten. Trends über eine einzige Sitzung oder über aufeinanderfolgende Sitzungen täglich oder wöchentlich wurden studiert.

Das zweite Experiment beabsichtigte, den Einfluss der " Zieldichte " (Zahl der gleichzeitig gezeigten Kontakte) und den Einflus der Wahrscheinlichkeit des Auftretens eines Kontaktes auf das Entdecken eines neuen Kontaktes festzustellen. Die experimentellen Sitzungen dauerten 30 Minuten. Zwischen den beiden Variablen wurde eine Wechselwirkung nachgewiesen.

Es wird über eine dritte Untersuchung berichtet, in der die Augenbewegungsmuster, die bei der Radarbeobachtung benutzt werden, durch elektro-oculographische Methoden registriert wurden. Dieses war nur eine Forschungsvorstudie, die aber gewichtige Beweise für ein un-

systematisches visuelles Prüfen brachte, mit grossen Unterschieden zwischen den Methoden, die von den Beobachtern benutzt wurden.

Die Gesamtergebnisse und ihre Tragweiten für das Radar-Personal werden diskutiert.

REFERENCES

ANDERSON, I. H., *et al.*, 1944, Radar operator fatigue: the effects of length and repetition of operating periods on efficiency of performance. United States O.S.R.D. Report No. 3334, Research Report No. 6, 1944.

BAKER, C. H., 1958, Attention to visual displays during a vigilance task: 1. Biassing attention *Brit. J. Psychol.*, **49,** 279–288.

BEXTON, W. H., HERON, W. and SCOTT, T. H., 1954, Effects of decreased variation in the sensory environment. *Canad. J. Psychol.*, **8,** 70–76.

BROADBENT, D. E., 1958, *Perception and communication* (Chap. 4), (Pergamon Press, London).

DEESE, J., and ORMOND, E., 1953. Studies of detectability during continuous visual search. U.S. Air Force, Wright-Patterson Air Force Base. W.A.D.C. Tech. Rept. 53–8.

ELLIOTT, E., 1957, Auditory vigilance tasks, Symposium on vigilance. *The Advancement of Science*, **53.**

MACKWORTH, N. H., 1950, *Researches on the measurement of human performance.* Medical Research Council Special Report Series No. 269, H.M.S.O.

MORGAN, C. T., 1952, Theory and problems of radar visibility: intensity modulated scopes, Johns Hopkins University and U.S. Naval Research Laboratory, Washington. N.R.L. Report 3965/JHU Report 166–1–148.

SAMUEL, J. A., 1958, A validation study of the use of the eye-movement measuring apparatus. *Admiralty Research Laboratory, Teddington. A.R.L./R.75/N.M.S.U.*

SHACKEL, B., SLOAN, R. C., and WARR, H. J. J., 1958 Detector plots eye-movements. *Electronics* **31,** 5.

SMITH, A. A. and BOYES, G. E., 1957, Visibility on radar screens; the effect of CRT bias and ambient illumination. *J. app. Psychol.*, **41,** 15–18.

WALLIS, D. and NEWTON, G. DE C., 1957, An experiment on synthetic signal injection as a method of improving radar detection. Admiralty, Division of the Senior Psychologist, *S.P. Report W2/4.*

WALLIS, D. and SAMUEL, J. A., 1956, Unpublished Admiralty Report. (Details of data breakdown and statistical analysis available on request from the authors.)

SOME CHARACTERISTICS OF INSTRUMENT MEASURING DIALS

By Antoine Papaloïzos
Department of Industrial Psychology, Ebauches, S.A., Neuchâtel, Switzerland

Two experiments were undertaken to evaluate the legibility of dials. The results of both experiments showed that one factor was extremely important for correct and quick reading : that the major graduations (the 10s and 5s) on the one hand and the dial + minor graduations on the other should be in a relationship of figure and ground. The special case of omitting minor graduations from the scale is also examined.

This paper reports two experiments which were undertaken to determine the importance of different characteristics of dials with respect to their legibility. The field of dial design has been reviewed recently by several authors (Chambers 1956, Murrell 1957, Tufts College 1952, Woodson 1956) and a large number of contributions have been made. The relevant variables are, however, many and there is still room for a great deal of further work.

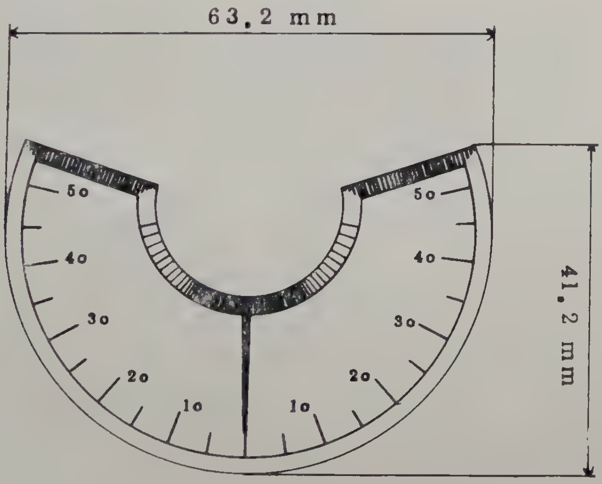
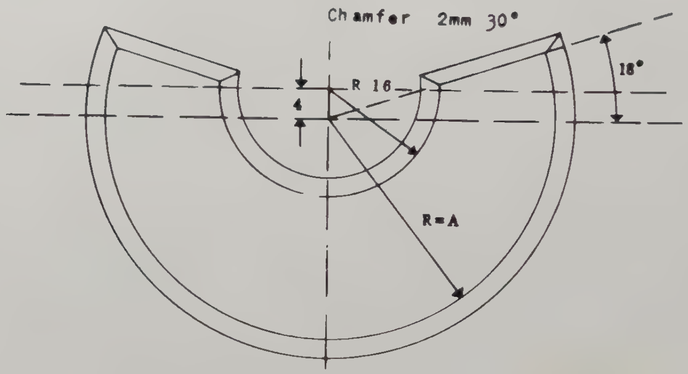


Figure 1. General view of the dial.

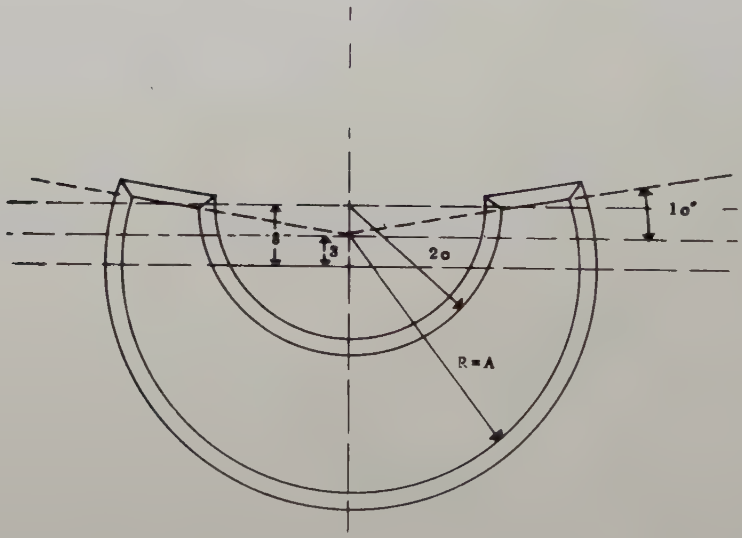
The dials with which the present study was concerned were to be fitted to a comparator of predetermined shape and dimensions shown in Fig. 1. These two factors were therefore given and could not be changed. Although in consequence restricted in scope, the study is believed to confirm and, in certain respects, to amplify previous knowledge relevant to dial design, especially as regards colour, upon which little previous work seems to have been done. The results suggest a principle of dial legibility which is discussed.

§ 1. Experiment I

Four areas were chosen for study : dial frame, dial hand, graduations and dial colour.



Small upper arc A=27.6mm



Large upper arc A=29.6mm

Figure 2.

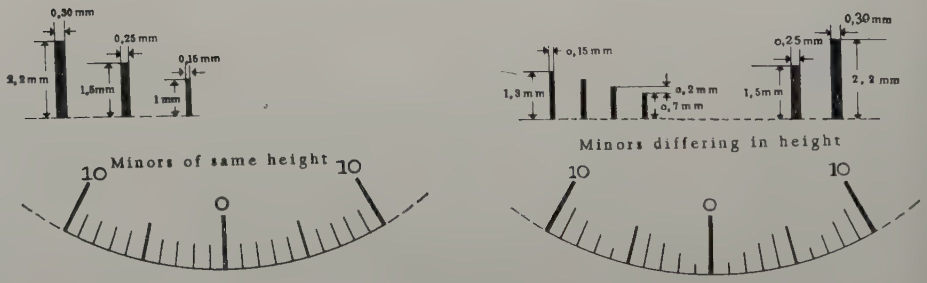


Figure 3.

For the first three areas, three variables, each with two levels, were investigated. For the fourth area (dial colour) there were two variables, one with two levels and one with four. Thus, for each area eight possible combinations were obtained, 2 × 2 × 2 for the first three and 2 × 4 for the last as shown in Table 1.

Table 1. Scheme of Experiment I

Area of study	Variable 1	Variable 2	Variable 3
1. DIAL FRAME	Saturation of grey :	Upper arc of window :	Finish :
	(*a*) Light (Munsell valve N7)	(*a*) Small	(*a*) Matt
	(*b*) Dark (Munsell valve N4)	(*b*) Large	(*b*) Glossy
		(The dimensions are shown in Fig. 2)	
2. DIAL HAND	Width of pointed end :	Length :	Colour :
	(*a*) Thick (i.e. the same width as the major graduations 0, 5, 10, 15, etc.)	(*a*) Long (i.e. overlapping the minor graduaitons)	(*a*) Same as major graduations
	(*b*) Thin (i.e. the same width as the minor graduations)	(*b*) Short (i.e. touching but not overlapping the minor graduations	(*b*) Different from major graduations
3. GRADUATION	Colours of majors and minors :	Colour of minors :	Height of minors (see Fig. 3) :
	(*a*) Same	(*a*) All one colour	(*a*) All the same
	(*b*) Different	(*b*) Two colours	(*b*) Differing
4. DIAL COLOUR	Saturation :	Hue :	
	(*a*) Light	(*a*) Neutral	
	(*b*) Dark	(*b*) Yellow	
		(*c*) Green	
		(*d*) Blue	

The colours used for the graduations and dials are set out in Table 2.

Eight subjects participated, 4 men and 4 women. Two of the men and two of the women had already worked with comparators ; the other four subjects had no experience of this type of work.

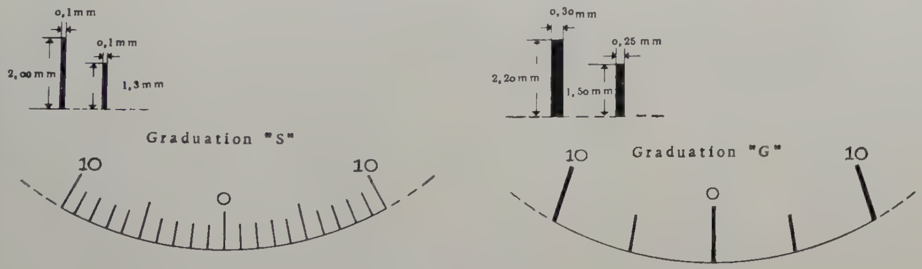


Figure 4.

64 dials were constructed, representing the combinations of the eight graduations with the eight colours. These dials were in the form of circular disks which were placed in a vertical position on a screen at a distance of 30 cm from the subject's eyes and under an illumination of 240 lx. The dial frame and the hand could be fixed on the apparatus separately. On the

experimenter's side this consisted of a scale on which a pointer could be moved to the position corresponding to the dial hand on the subject's side. A shutter controlled by an electric timer enabled exposures of 0·2 sec to be made.

A factorial design randomised the following variables: subjects, dial frame, dial hand, graduations, colour of dial, time of day when readings were made, day on which readings were made, lists of readings. There were eight lists of numbers which constituted the readings to be made. The numbers

TABLE 2. COLOURS USED IN EXPERIMENT I

Dial No.	Dial hue	Munsell notation	Hue of major graduations	Length of minor graduation	Hue of minor minor graduations below "5" s	Munsell notation	Hue of minor minor graduations above "5" s	Sign – and –
1	white	N9	black	same length	strong blue	5PB 3/10	strong blue	b bottom
2	,,		,,	,,	cherry red	7.5R/14	black	b top
3	,,		,,	,,	black	N1	black	b ,,
4	,,		,,	,,	cherry red	a	strong blue	b bottom
5	,,		,,	different length	black	a	black	b ,,
6	,,		,,	,,	cherry red	a	strong blue	b top
7	,,		,,	,,	strong blue	a	strong blue	b ,,
8	,,		,,	,,	cherry red	a	black	b bottom
9	pale yellow	2.5Y 8/4	black N1	same length	strong blue	a	strong blue	b bottom
10	,,		,,	,,	cherry red	a	black	b top
11	,,		,,	,,	black	a	black	b ,,
12	,,		,,	,,	cherry red	a	strong blue	b bottom
13	,,		,,	different length	black	a	black	b ,,
14	,,		,,	,,	cherry red	a	strong blue	b top
15	,,		,,	,,	strong blue	a	strong blue	b ,,
16	,,		,,	,,	cherry red	a	black	b bottom
17	grayish yellow	5Y 7/4	white N9	same length	dark green	2.59 3/6	dark green	b bottom
18	,,		,,	,,	cherry red	a	white	b top
19	,,		,,	,,	white	a	white	b ,,
20	,,		,,	,,	cherry red	a	dark green	b bottom
21	,,		,,	different length	white	a	white	b ,,
22	,,		,,	,,	cherry red	a	dark green	b top
23	,,		,,	,,	dark green	a	dark green	b ,,
24	,,		,,	,,	cherry red	a	white	b bottom
25	light blue	10 G 8/2	black N1	same length	dark green	a	dark green	b bottom
26	,,		,,	,,	cherry red	a	black	b top
27	,,		,,	,,	black	a	black	b ,,
28	,,		,,	,,	cherry red	a	dark green	b bottom
29	,,		,,	different length	black	a	black	b ,,
30	,,		,,	,,	cherry red	a	dark green	b top
31	,,		,,	,,	dark green	a	dark green	b ,,
32	,,		,,	,,	cherry red	a	black	b bottom
33	light greenish blue	7.5B.G7/6	black N1	same length	chinese orange	7.5R6/12	chinese orange	bottom
34	,,		,,	,,	cherry red	a	black	b top
35	,,		,,	,,	black	a	black	b ,,
36	,,		,,	,,	cherry red	a	chinese orange	bottom
37	,,		,,	different length	black	a	black	b ,,
38	,,		,,	,,	cherry red	a	chinese orange	top
39	,,		,,	,,	chinese orange	a	chinese orange	,,
40	,,		,,	,,	cherry red	a	black	b bottom
41	dark greenish blue	2.5B 2/2	white N9	same length	chinese orange	a	chinese orange	bottom
42	,,		,,	,,	pale blue white	2.5B 7/4	white	b top
43	,,		,,	,,	white	a	white	b ,,
44	,,		,,	,,	pale blue white	a	chinese orange	bottom
45	,,		,,	different length	white	a	white	b ,,
46	,,		,,	,,	pale blue white	a	chinese orange	top
47	,,		,,	,,	chinese orange	a	chinese orange	,,
48	,,		,,	,,	pale blue white	a	white	b bottom
49	dark blue	5.PB 2/2	white N9	same length	strong yellow	5.Y 8/10	strong yellow	b bottom
50	,,		,,	,,	strong blue	a	white	b top
51	,,		,,	,,	white	a	white	b ,,
52	,,		,,	,,	strong blue	a	strong yellow	b bottom
53	,,		,,	different length	white	a	white	b ,,
54	,,		,,	,,	strong blue	a	strong yellow	b top
55	,,		,,	,,	strong yellow	a	strong yellow	b ,,
56	,,		,,	,,	strong blue	a	white	b bottom
57	black	N1	white N9	same length	strong yellow	a	strong yellow	b bottom
58	,,		,,	,,	strong blue	a	white	b top
59	,,		,,	,,	white	a	white	b ,,
60	,,		,,	,,	strong blue	a	strong yellow	b bottom
61	,,		,,	different length	white	a	white	b ,,
62	,,		,,	,,	strong blue	a	strong yellow	b top
63	,,		,,	,,	strong yellow	a	strong yellow	b ,,
64	,,		,,	,,	strong blue	a	white	b bottom

a = Munsell notations appear once for the same colour. b = Munsell notations same as preceding column.

were drawn from a table of random numbers normally distributed with a mean of zero and a standard deviation of 10. Each subject came once every day for eight consecutive working days and made 200 readings at each sitting. These gave a total of 12 800 readings.

Parallel to the main investigation, two further types of graduation, shown in Fig. 4, were studied. The first (S) was a standard type currently used in comparators ; the widths of its major and minor graduations are the same but the majors are longer than the minors.

The second type of graduation (G) had no minors. Both (' S ') and (' G ') appeared in the same colours as the other graduations and each subject made 100 readings with each type and colour. The resulting 800 readings with (' S ') and 800 with '(G ') were randomly attributed, in sets of 100 readings, to the eight testing days.

1.1. *Results*

The subjects differed widely in the extent to which their performances improved over the eight days. Accordingly, a regression coefficient was calculated for each subject and a *deviation score* was obtained by subtracting the predicted score from the actual score. The deviation scores were used for analyses of variance.

The experimental design permitted the analysis of nine sources of variation. The first four were the areas of investigation and the other five were the subjects, the day, the time of day, the lists of numbers and the variation between the sessions with 200 and 300 readings.

The total number of errors was 2542 over 12 800 readings. Thus about 20 per cent of the readings were errors.

Dial frame

Results were found not to differ significantly for any of the three variables studied, although there was a significant interaction between saturation and finish ($p < 0{\cdot}05$), indicating that a light glossy grey or a dark matt grey (1210 errors) were better than a light matt grey or a dark glossy grey (1332 errors).

Dial hand. Again results were found not to differ significantly for any of the three variables studied. There was, however, a significant interaction between Length and Colour ($p0 < {\cdot}05$), indicating that a long hand in the same colour as the major graduations and a ' short ' hand in a different colour from them (1205 errors) were better than a ' long ' hand in a different colour from the major graduations and a ' short ' hand in the same colour as them (1337 errors).

Graduations

Results for the first variable were significant ($p < 0{\cdot}05$), indicating that a different colour for the major and minor graduations was better than the same colour for both types (1202 errors v/s 1340).

Dial colour

Again results for the first variable were significant ($p < 0{\cdot}01$), indicating that a light colour (1194 errors) was better than a dark one (1348 errors).

The interaction of saturation and colour was also significant ($p < 0{\cdot}01$) indicating that the variance between the eight colours was significant. The percentages of errors for the eight colours were as follows:

	per cent		per cent.
dark green	17	light green	21
pale yellow	17	dark blue	21
light blue	18	black	22
white	19	greyish yellow	25

A percentage of 2·6 is necessary for a difference to be significant at the 0·5 level.

It should be pointed out that the experimental design permitted the analysis of results only within each area. The interaction between, say, graduations and dial hand is not therefore statistically possible. What can be done, however, is to combine the best results obtained *within* each area and thus reconstitute a whole comparator.

Results with graduations ' S ' and ' G '. The following findings are based on the original scores (errors) and not on the deviation scores. This is due to the fact that no learning was possible with ' S ' and ' G ', since these two types of graduation were presented only once to each subject. The percentage of errors made with the eight graduations taken together is 25 (the highest being 28 and the lowest 22). The percentage made with ' S ' was 41, while that made with ' G ' was 17. The superiority of ' G ' is in line with results by Loucks (1944).

The figures seem to indicate that ' S ' and ' G ' compared to the other eight types of graduations the worst and the best, respectively. To what extent, however, are these results due to chance ? To answer this question, the Wilcoxon matched-pairs signed ranks test (cf. Siegel 1956) was applied to the data as set out in Table 3. When ' S ' is compared to the eight graduations, Wilcoxon's $T = 1$ which with $N = 8$ has a probability of 0·02 (2 tail-test). Thus the results with ' S ' can occur by chance twice in a hundred times. When ' G ' is compared to the eight graduations, Wilcoxon's $T = 4{\cdot}5$. With $N = 8$, Wilcoxon's table, at the chance level of 0·05 (2 tail test) shows a T of 4. Thus the chance probability attached to the obtained T of 4·5 remains indeterminate, but is in the region of 0·05. The sign-test (cf. Siegel) applied to the same data, shows that the chance probability of obtaining for the ' G ' results when compared to the eight graduations results is $p = 0{\cdot}07$ (2 tail test). If only the 0·05 chance level is considered as being ' significant ' then neither the Wilcoxon test nor the sign test indicate a ' significant ' result for the ' G ' graduation when compared to the eight other types taken together.

Table 3. Percentages of errors made with different types of graduation in Experiment I

	Subjects							
	1	2	3	4	5	6	7	8
8 graduations together	26	29	29	34	38	16	14	16
graduation ' S '	58	41	47	43	50	15	37	36
graduation ' G '	19	19	20	13	28	8	23	8

Other variables. Each day was divided into eight hours, one for each of the eight subjects. The number of errors made in the morning as against those made in the afternoon (summed over subjects and over days) was not significant. A second variable was the number of errors made at the beginning of the half-day (first two hours a.m. and first two hours p.m.) compared with those made at the end of the half-day (last two hours a.m. and last two p.m.). This variable was significant at the 0·01 level, the errors being more numerous at the beginning than at the end of the half-days.

Variation between results for different lists of numbers was significant at the 0·05 level. The variation between numbers of errors made in sessions with 200 and 300 readings was not significant.

1. *Nature of errors made.* Number of errors made on the right (positive) and left (negative) sides of the scales The errors were distributed in the following way for each of the eight subjects.

Subject	1 (Male)	2 (M)	3 (M)	4 (M)	5 (fem)	6 (F)	7 (F)	8 (F)	Total
Number of errors (right side)	352	138	299	257	161	239	130	225	1801
Number of errors (left side)	302	129	245	247	152	267	156	246	1744

A binomial test applied to the errors made on the right and left sides (1801 and 1744), resulted in a $z = 0{\cdot}96$ with $p = 0{\cdot}17$. Thus the distribution of errors on each side is not significantly different from a chance distribution.

2. *Systematic errors of* 5 *or* 10 *points.* Were the errors made on each of the different types of graduation of a special nature ? For instance, was there a tendency for errors of 5 or 10 points (plus or minus) to be made more frequently on one type of graduation than on another ? Such errors were counted and more than 80 per cent were in the 5 points category, the remainder in the 10 points ; no 15 points errors were found. These errors were grouped according to the eight subjects and the 10 graduations (8×10 table) and Friedman's two-way analysis of variance was applied (cf. Siegel 1956). The result is a $\chi_r^2 = 2{\cdot}88$ with 9 degrees of freedom and the chance level is 0·95. We conclude that the errors of 5 or 10 points made on the 10 graduations were not different from what one would expect by chance.

§ 1.2 *Discussion*

The striking differences found between the three types of graduations (' S ', ' G ', and the other eight,) suggest that the main information for correct dial reading is given by the major graduations ; the minors ' inform ' only if they do not ' mask ' the information given by the majors. Thus the best graduation was that of ' G ' with no minors. The next best were those of the eight main types in which majors and minors were distinct from each other in weight and height, while the distinction in ' S ', the worst type, was not at all clear, the width being the same and only the 10s differing in length. This principle is consistent with much previous work. Do the other significant findings of the present experiment conform with it ?

The finding that it was preferable to have majors and minors of two different colours rather than the same colour, fits very well and does not need any comment. The finding that a light colour dial tended to be preferable to a dark one is of doubtful relevance. The major graduations were always either white or black, white when the dial colour was dark and black when the dial was light. Thus, there was always a contrast between the colour of the majors and the colour of the dial which formed the background ; it could, however, be argued that a black figure on a light background stands out better than a white figure on a dark background, since material in everyday use, such as printing, is of the former kind. If so, the significant finding about saturation of dial colour can perhaps fit in with the principle. The three significant interactions in the Dial Frame, Saturation and Hand areas do not fit in with the principle, but they do not contradict it either. They seem to be independent.

The principle is thus fairly well supported as far as the present data go. There are, however, a number of possible objections due to the limits of the experimental design, of which two may be especially mentioned. The first is the short exposure time of the stimuli : are the results obtained a function of this experimental condition ? The second objection is that in readings presented to the subjects, the dial hand was always exactly on a unit and never between two units. Both these objections have some force. Although it could be argued that the dial which has given proof of its comparative advantages under these conditions will also show its superiority under more ‘ normal ’ ones, discrepancies between the findings of previous investigations indicates the need for caution. The second experiment was designed to overcome both of these objections and at the same time to cover a greater variety of graduations.

§ 2. *Experiment II*

In the second experiment 16 different dials were used, a dial being a combination of one of eight different types of graduation and one of 2 types of hand. Fig. 5 shows each of these types and the symbols used to denote them. The dials are shown in Fig. 5 at approximately full size.

Sixteen subjects, eight women and eight men, participated ; each came twice and at each sitting made 160 readings, 10 for each of the 16 dials. These readings were random numbers with one decimal and the subject had to round off to a unit (e.g. 10·4 was 10 ; 10·6 was 11 ; there were no 0·5s). The same 160 readings were used for each subject so that the difficulty was uniform.

The order in which the 16 dials were presented differed for each subject. A latin square of 16 × 16 was used, each subject having the order of presentation of a row : thus, over the 16 subjects each dial preceded and followed each one of the other dials. In the second sitting for each subject the dials were presented in the reverse order.

The dials were shown to the subjects by means of a film. The subject, seated in front of a screen, pressed a lever opening a small shuttter in the middle of the screen on which a dial was projected. The subject was instructed to take all the time necessary for the correct reading, then call out the number and depress the lever which he still held. This closed the shutter ; the opening and shutting times (in 1/100 sec) were automatically recorded on paper ; if a mistake was made in the reading this was recorded by the experimenter. A study by Elkin (1959) has shown that the time to verbalise is a constant, regardless of difficulty in reading. We therefore have two criteria : time taken to make a reading and number of errors.

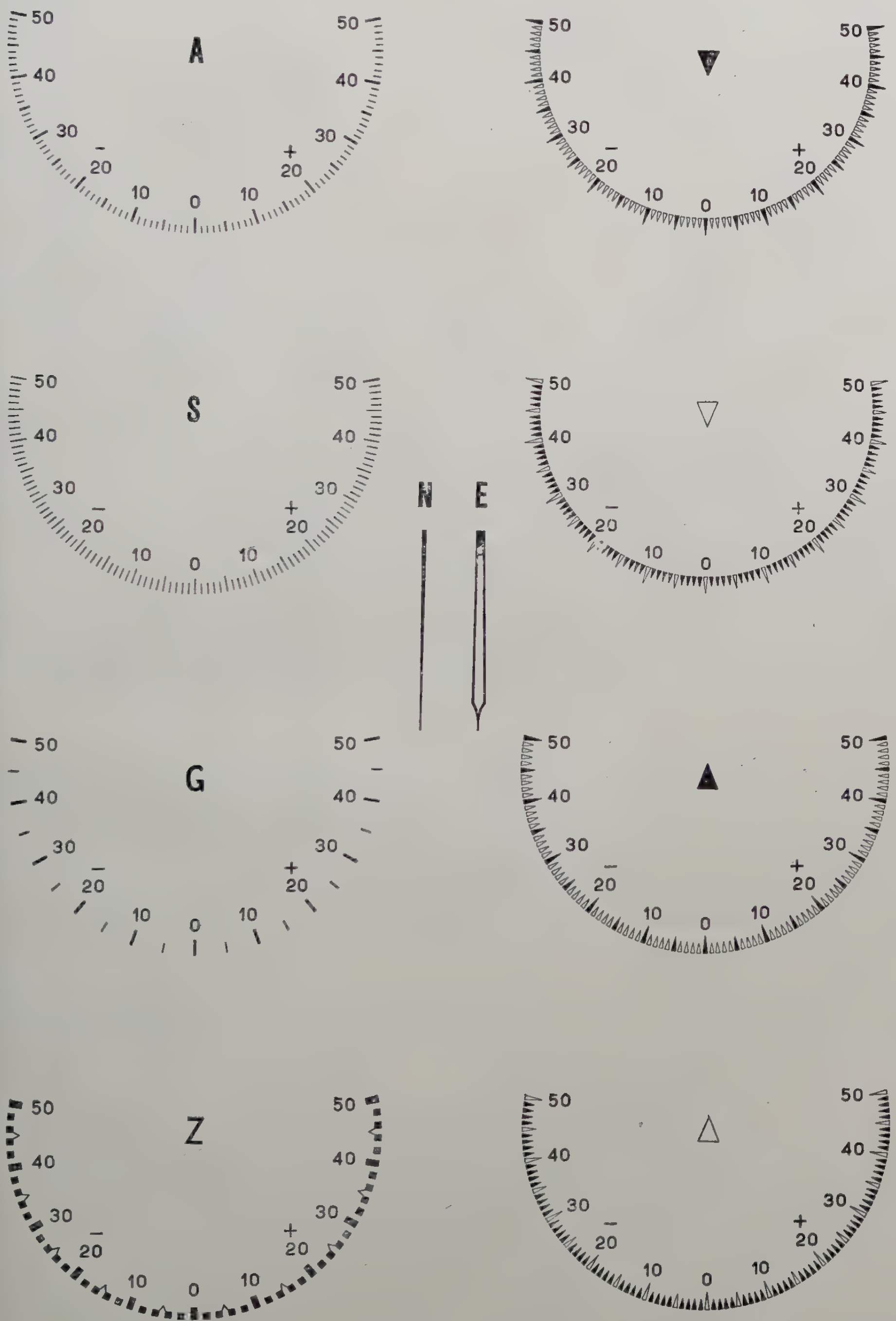


Figure 5. Dial and dial hands used in Experiment II.

§ 2.1 *Results*

Times. The distribution of the reading times was not normal, as is frequently the case with time measurements. A logarithmic transformation was therefore used and the resulting scores were subjected to an analysis of variance the results of which are shown in Table 4. The F obtained for the variation between dials being highly significant, *t*-tests were carried out to compare performances with the individual dials.

Table 4. Experiment II. Analysis of variance of log times

Source of variation	degrees of freedom	Mean square	F	P
Subjects (S)	15	5·6859	302·44	0·01
Dials (D)	15	0·8911	47·40	0·01
Ciphers in groups of 10 (C)	19	0·4363	23·21	0·01
Interactions				
S×D	225	0·0503	2·68	0·01
S×C	285	0·0215	1·14	n.s.
D×C	285	0·0151	0·80	n.s.
Error	4275	0·0188		
Total	5119	0·4082		

The mean times for the various dials are given in Table 5. The 16 dials were divided into two classes, one with the normal dial hand and the other with the wider hand. The difference between these two classes was not significant. The 16 dials were accordingly grouped in the eight graduation categories and *t*-tests were applied. Table 6 shows the results obtained. The eight dials can be grouped into four categories, the first two with three dials each and the last two one each as shown by the vertical lines in Table 6. The differences between dials in the same category are not significant but are so between one category and another (except for ‘ S ’ and ‘ G ’ in groups 2 and 3).

Errors. Table 7 shows the number and percent of errors made on each of the eight graduations. Friedman's 2-way analysis of variance was applied to the data cast in a 16×8 table (subjects and graduations) and the $\chi_r^2 = 270$, which shows a chance probability of below $< 0{\cdot}01$. This test tells us that the differences in numbers of errors are not due to chance, but does not specify the amount of difference needed between two graduations, which can be considered due to chance.

The Wilcoxon test (cf. Siegel 1956) was applied to pairs of graduations and the resulting Ts showed the following graduations as being different at the 0·05 level (from right to left in table 7) :

Z and G
G and S
S and △ (S and ▲ not being different)

There were no significant differences between the first five graduations.

We could thus form four groups of graduations :

▼ A ▽ △ ▲
▲ S
G
Z

Table 5. Experiment II. Mean times

	Geometric mean time in 1/100 sec															
	162·6	164·8	164·9	166·4	166·8	170 6	176·2	178·3	178·6	180·3	180·8	184·1	188·1	193·2	240·5	244·2
Type of graduation	A	▼	▼	▲	A	▲	▽	S	△	△	▽	G	S	G	Z	Z
Dial hand	E	N	E	N	N	E	N	N	N	E	E	N	E	E	E	N

A difference of 5/100 sec or more is needed for t to be significant at the 0·5 level

Table 6. Experiment II. Grouped mean times

Geometric mean time in 1/100 sec	164·7	164·8	168·5	178·5	179·5	183·2	188·6	241·8
Type of graduation	A	▼	▲	▽	△	S	G	Z

A difference of 6/100 sec for t to be significant at the 0·05 level

Table 7. Experiment II. Errors

Number of errors	20	22	23	25	30	41	76	122
Percent of errors	3·1	3·4	3·6	3·9	4·7	6·4	11·9	19·1
Type of graduation	▼	A	▽	△	▲	S	G	Z

A difference of 15 errors or more is needed for χ^2 to be significant at the 0·05 level

Each of these groups is significantly different from its adjacent groups except that ▲ is not different from either the first or the second group and has therefore been listed in both.

Relation between errors and time. A rank correlation coefficient (Spearman's rho) between the number of errors and time taken was computed for the 16 dials. A highly significant (0·01 level) positive correlation of 0·79 was obtained i.e. the dials with a short reading time had also few errors.

2.2. *Discussion*

At the end of the first experiment we stated that the major graduations provided the main information in dial reading and that the information provided by the minor graduations should not ‘ mask ’ that of the major. Are the results of the second experiment compatible with this statement ?

Let us consider the groupings of Table 6 (those of Table 7 are quite parallel). The main difference between groups 1 and 2 lies in the fact that in the former the major graduation is black while in the latter it is white (▲ and ▼ vs and △ and ▽) ; the background, for all graduations, being a white semi-transparent paper, the black graduation will be more apparent than the white one. In groups 1 and 2 we also find ‘ A ’ and ‘ S ’. These two graduations are also differentiated by the width of the majors (cf. Fig. 5). The majors in group 1 are thus in all cases more prominent than the minors, while in group 2 they are either less prominent or very little different from the minors. Thus the first 2 groups fit in perfectly with our statement above. In the third group we find ‘ G ’ which was the best type of graduation in the first experiment. ‘ G ’ has no minor graduations, so that our statement about the information of majors and minors is not supported, since it would have predicted a better result for ‘ G ’.

The following table presents the results with ‘ A ’ and ‘ G ’ in the two experiments.

	First experiment		Second experiment	
	Time of reading	Errors per cent	Time of reading	Errors per cent
A	0·2 sec	22	1·65 sec	3
G	0·2 sec	17	1·89 sec	12

When ample time is given, there is a reduction in the number of errors made for both ‘ A ’ and ‘ G ’, but this reduction is considerably greater for ‘ A ’ than for ’ G ’. The relationship between reading speed and number of errors for ‘ G ’ is worth further experimentation. With very quick readings (0·2 sec) ‘ G ’ is the best type of the graduations we studied ; but when the subject takes his time (1·5 sec or more) ‘ G ’ is one of the worst types. This finding is in line with those of Kappauf, Smith and Bray (1947) and Kappauf and Smith (1948). What happens to ‘ G ’ when the reading speed is between 0·2 and 1·5 of a second ? Does the percentage of errors made with ‘ G ’ fall below 12 per cent at any point ? In other words is the relation between speed and errors linear or curvilinear ? This question may be important when the situation calls for very quick readings to be made.

The results with graduation ‘ Z ’ (worst type) fits in with our statement about the information of majors and minors since they were in this type relatively poorly differentiated. The results with 'Z' should, however, be taken with some caution, as it was quite a different type of graduation from the others studied, demanding a different perceptual set for its reading; furthermore a fault in its design was observed during testing: when the dial hand was on the separation between two zones (i.e. on 0·4 or 0·6) it extended the black zone so that the reading was quite difficult. In readings involving 0·4 or 0·6 the hand should have overlapped the zone's height by a fraction only.

§ 3. CONCLUSIONS

Results of both experiments can be summed up by saying that for correct and quick dial reading the relationship between the major graduations on the one hand and the minor graduations and dial on the other, should be one of 'figure' and 'ground'. By a figure-ground relationship we mean that the major graduations must clearly stand out of, and be distinct from, the rest of the field of perception. Although what is 'figure' and what is 'ground' are to some extent under the voluntary control of the observer, by focussing the attention on one part or another of the field of view, there are certain physical characteristics, such as boldness of lines, colour, contrast, etc., and organisation of the whole, which make some parts more prominent, more 'figural' than others. In the case of dials, given two types of graduation, one having lines thicker or bolder in some other way than the other, and the same dial colour for both, the graduation with the bolder lines will be more prominent, more 'figural' than the other.

The special case of omitting altogether the minor graduations, is advantageous for very quick readings (0·2 sec) but not with ample time (1·5 sec). In between these two reading speeds, the special case is still indeterminate.

We wish to express our thanks to Mr. Jean Cardinet for the experimental design, to Mr. Ferdinand Maire for preparing the dials and to Mr. Roger Kramer for making the apparatus.

Le présent article relate deux expériences entreprises pour apprécier la lisibilité de cadrans destinés à équiper des appareils de mesure. Les résultats des deux expériences montrent qu'un facteur favorise une lecture rapide et correcte. Ce facteur stipule que les graduations principales (10 et 5) d'une part, et, d'autre part, le cadran et les graduations secondaires, se présentent dans le même rapport qu'une figure sur un fond.

Zwei Experimente wurden ausgeführt zur Schätzung der Lesbarkeit von Zifferblättern. Die Resultate beider Experimente zeigten die Wichtigkeit eines Faktors für fehlerfreies und schnelles Ablesen: nähmlich dass die Hauptunterteilungen (für 10 und 5) einerseits und das Zifferblatt und die kleineren Unterteilungen anderseits in Verhältnis zu Figur und Grund sein sollen.

REFERENCES

CHAMBERS, E. G., 1956, Designs and uses of scales, dials and indicators. *Applied Psychology Research Unit, Cambridge, England.*

ELKIN, E. H., 1959, Effects of scale shape, exposure time and display-response complexity on scale reading efficiency. Wright Patterson Air Force Base, Ohio, *USA-WADC Technical* Report 58.472.

KAPPAUF, W. E., SMITH, W. M., and BRAY, C. W., 1947, Design of instrument dials for maximum legibility. I. Development of methodology and some preliminary results. *USAF Air Materiel Command Memo.* Report No. TSEAA-694-IL.

KAPPAUF, W. E. and SMITH, W. M., 1948, Design of instrument dials for maximum legibility. II. A preliminary experiment on dial size and graduation. *USAF Air Materiel Command Memo.* Report No. MCREXD-IN-694.

LOUCKS, R. B., 1944, Legibility of aircraft instrument dials : the relative Legibility of tachometer dials. *AFF School of Aviation Medicine Project* No. 265 Report No. 1.

MURRELL, K. F. H., 1957, Data on human performance for engineering designers. *Engineering*, pp. 1–32.

SIEGEL, S., 1956, Nonparametric statistics for the behaviour sciences (New York : McGraw-Hill).

TUFTS COLLEGE INSTITUTE OF APPLIED EXPERIMENTAL PSYCHOLOGY, 1952, *Handbook of human engineering data*, 2nd ed. *Navexos* P-643.

WOODSON, W. E., 1956. *Human Engineering Guide for Equipment Designers* (University of California Press).

DIFFERENTIAL EFFECTS OF KNOWLEDGE OF RESULTS AND DISCUSSION

By B. T. Jensen, J. R. Tilton and D. N. Anderson

System Development Corporation
Santa Monica, California

A training experiment was initiated at a radar station of the United States Air Force. External and unanticipated events so attenuated the experimental conditions that the effects cannot be evaluated unambiguously. However, the change in performance of crews supports the notion that knowledge of results and discussions contribute to improvement by helping to develop understanding of the environment and common modes of response among members of a group. The paper describes the study and details the changes in performance in relation to the kinds of information available to the crews.

§ 1. Introduction

For a number of years the United States Air Force Air Defense Command (ADC) has had a system training programme designed to improve the day-to-day proficiency of radar crews and to increase their flexibility in responding to new situations which might arise in the event of hostilities. As part of the programme, observers of a specific training exercise provide knowledge of results of the crew's operations by reporting to the crew members regarding their achievement; the crew then discusses its work and procedures in an attempt to improve operations. Although this combination has been demonstrably effective, questions were raised as to the relative contributions of the knowledge of results and of the discussion sessions. An attempt was made to gain information in a field setting by experimentally manipulating the training programme in an operational radar station. Conditions beyond the experimenter's control weakened the study, but it nevertheless yielded valuable observational data.

The data do not lend themselves to refined statistical analysis, but do indicate that the report to the crews and the discussions had strong effects. Specifically, they aided in developing, among the crew members, a common agreement regarding responses and a common and correct 'model of the environment', in the sense of an understanding of requirements and the responsiveness of the external world to crew activities.

The research was conducted in a military setting using ADC's System Training Programme (STP). The study involved only four radar crews. However, we believe that intensive analysis of a small number of cases, when related to a conceptual framework based upon study of other cases, does lead to significant interpretation. The comments made in this paper draw upon considerable experience of the writers in developing programmes for, and monitoring and doing research upon, the training of man-machine systems both prior to and following the experiment reported here.

§ 2. The Setting

The STP was based upon experimentation in the System Research Laboratory of the RAND Corporation. There, subjects organized as crews, were presented with simulated radar information indicating all air traffic in a given area (some of it supposedly hostile) and the crews would attempt to defend the area by performing activities leading to the shooting down of the invaders. Immediately after this type of excerise, observers reported to the crew on its achievement relative to the stimulus; e.g., ‘ The crew intercepted ten of twelve invader aircraft.’ The crews participated in problem-solving discussions after receiving the reports. Observers noted a marked improvement in the performance of the crews working under some of the experimental conditions. Further experiments and development work led to STP in which radar crews of the ADC worked at their tasks of responding to simulated air traffic appearing on radar scopes during a two-hour exercise which was followed by performance information (called ‘ feedback ’), and discussions (called ‘ debriefings ’). Additional information regarding the programme and its development can be found in papers by Alexander, *et al.* (1959), Chapman, *et al.* (1959) and Goodwin (1957).

Continued experimentation indicated ways in which feedback should be changed with subsequent gain in achievement by the crew in training. A field experiment utilizing real crews at a temporarily inactive radar station indicated clear-cut superiority of crews having the new feedback and debriefings as compared to those experiencing only the training exercises.

§ 3. Plan of the Present Study

The present study examined performance of crews participating in the training exercise under four sets of conditions designed to enable feedback and debriefing to be studied separately and in combination as shown in Table 1.

		Debriefing	
		Yes	No
Feedback	Yes	Crew FD	Crew F
	No	Crew D	Crew O

Table 1. Experimental Conditions

The observations were made at a newly-constructed radar station on the California coast. The ADC had agreed to continue this particular station on an inactive status (subject to changing defence needs) for about three months to permit experimental modification of the usual programme. The Commander of the station assigned thirteen enlisted men with one or more years of Air Force experience to serve as training personnel and to collect data required. One or more civilian representatives of the System Development Corporation was at the station at all times during the conduct of STP activities.

Standard materials in STP include a problem film which, when fed through a special type projector, creates a dynamic air picture on the radar scopes. To aid in preparing feedback, the observer is provided with a set of maps giving the position of each aircraft at each five minutes throughout the two-hour exercise, scripts for simulated inputs and recording forms which list

' critical ' flights (those of special concern in that exercise) and describe characteristics of those aircraft. For this study, special films had been prepared giving extra heavy traffic in the area to be defended by this station. The various films included a peacetime problem with situations requiring special action; for instance, an air-sea rescue mission looking for survivors. Other problems included wartime situations with ' invader ' aircraft, and SAC flights on retaliation missions. Scripts were prepared for the simulators of adjacent stations indicating what information should be passed into the station from each of several agencies, and at what times. Since not all actions and decisions could be standardized in script form, a guide was prepared for the simulator of higher headquarters indicating what action he should take in regard to different items of information which might be passed to him.

§ 4. Sequence of Events

This section describes what actually happened, not what was planned. From a pool of enlisted men and junior officers, four radar crews were organized so that all were approximately equal in rank and experience of the members. Most of the members had no prior experience in operating crews. Each crew was introduced to the training programme by lectures covering the purpose and the nature of the programmes, and the geographical area of responsibility–that is, the areas to be defended by this crew and by crews at adjacent radar stations. The lectures were followed by two practice, or shakedown, sessions to permit each crew to effect its internal organization and to become familiar with the conduct of STP.

Following these each crew was given a rather heavy-load wartime problem with many aircraft on the radar scope, a large proportion of which were invaders This exercise served as a pre-test and was repeated three months later.

Within two hours after the pre-test exercise for the fourth crew, the Air Defense Command activated the station. It was then manned around the clock, so that the crews which should have been trained in the experimental programme were experiencing at least forty hours of live operations each week. Furthermore, a number of new men, previously unassigned to operations, were placed in the crews for training purposes; however, men did not serve on more than one crew except one officer and one crew chief (enlisted man) who were transferred shortly after the declaration of emergency.

After about ten days, ADC decided to continue the experimental training programme modified so as to permit continuous operations. This meant that during each training exercise, part of the crew dealt with live traffic and handled emergencies, and that on some occasions, STP was delayed or postponed because of operational demands.

During the next few weeks all crews experienced approximately the same number of STP exercises, receiving feedback or debriefing as planned (see Table 1). However, we know that crews had meetings during regular shifts, particularly on the midnight turn, at which they discussed operations, and no attempt was made to stop these activities. Furthermore, because of operational needs the regularly scheduled debriefings were many times cut short and were often not attended by one or more crew members. Another factor limiting the amount of time spent in debriefings was the fact that the training personnel manned the station during debriefing time and applied pressure upon the crews to return, particularly in the evenings when such relief indicated the end of the day for the training people.

§ 5. Discussion of Results

Because of equipment difficulties, which interrupted the exercise, and because four of the most capable persons were re-assigned on the morning of the post-test, the performance of Crew FD on the post-test was very poor—it was obvious to all that it was performing far below its usual level. For this reason, objective data on this crew have not been considered, although other observations of it contribute to the general discussion and interpretation of results.

The main changes between pre- and post-tests in the performances of the three remaining crews are set out in Table 2. Clearly the crews changed to different extents in the several component functions of their task. In the following paragraphs we attempt to illustrate how different understanding of the task led to the preoccupation of crews with different functions, and how feed back affects perception of the task. All crews had been informed that the primary mission of this station was the detection and tracking of aircraft appearing in certain parts of the area visible by radar. In as much as all started at about the same level of skill, we could expect all crews to gain an equal amount from practice. Deviation from uniformity may therefore be reasonably attributed to the effects of the differing treatments received by the crews.

Function	Crew F	Crew D	Crew O
Detection			
Per Cent	*	†	‡
Delay	*	†	‡
Tracking	†	†	†
Telling :			
Required	†	‡	*
Not Required	‡	*	‡

* No change
† Performance on Post-test more desirable
‡ Performance on post-test less desirable

Table 2. Relative Improvement of Three Crews

(a) *Detection.* The only crew to improve in this function was Crew D, who had debriefing time to discuss operations. Fortunately for the crew, the approach of the crew leaders was, " What problems do we have which keep us from doing our job ? " The discussions were aimed then at clarification of tasks, interpretation of regulations, and reducing the communication difficulties among members. Out of this kind of discussion one would expect a common understanding of goals and procedures to emerge. This was important because in all crews several persons performed the detection function, rotating in the position at the radar scopes; and at all times, at least two individuals observed different parts of the area simultaneously. The deterioration of Crew O at detection will be considered later.

(b) *Tracking*. All three crews improved about the same on the function of tracking after detection. Airborne objects reported by the observer were plotted with time of entry on a large plexiglass map in front of the room. All crew members could see this display; discrepancies and lack of currency were visible. One of the crew leaders had the specific responsibility of seeing that this display was current and accurate. It would be surprising if new crews did not improve in this highly visible task. Only one crew had feedback on tracking, but apparently the feedback contributed little under these conditions. (Papers by Berkowitz (1958) and by Geddes *et al.* (1959) provide data and comments regarding effects of making performance of a function visible.)

(c) *Telling*. The effects of information about the job are seen in the function of telling (reporting air traffic to other stations). Table 2 shows that Crew F improved in telling required information. This crew it will be remembered, received post-exercise reports on how many tracks it reported compared with the number that should have been told: in short, it was confronted regularly by evidence of inadequacies.

Crew F also however plotted a great deal of traffic in its ‘ free ’ area (i.e. safe areas that did not need to be watched routinely). The crew worked very hard to plot and maintain traffic, but it did so at the expense of maintaining good defence; its energies were expended in unnecessary ways because no one in the crew really understood the nature of the world about them. This was probably due to the fact that the feedback included the information that the crew was making unnecessary reports but gave no information about which kind of traffic should be reported to other stations.

Crew D, with no reminders about telling, became less efficient in performing several jobs requiring cooperation with other stations. Its attention seemed to be focussed on doing its more visible tasks.

Crew O, on the other hand, with neither feedback nor debriefing did not deteriorate in performing *required* telling. This difference from Crew D was due to the fact that Crew O did more telling than any other, although most of it was not required. This crew elected to do its task by ‘ brute force ’ ; the crew chief said to an experimenter, “ We don’t know what we are supposed to do, so we do everything we can.” The chief of Crew O had worked with Crew F for one week during which time the crew had one training exercise with feedback which included a report on telling. Thus, this crew chief knew that telling was important but he did not remember the requirements, even though they had been stated to all crews at the beginning of the study. A similar situation existed with respect to detection; Crew O detected and plotted a great deal of air traffic but most of it was in the so-called free area.

Looking at the data of Table 2 in terms of crews rather than functions, it can be seen that Crew O, having neither feedback nor debriefing deteriorated in both Detection and Telling. Crew D, with debriefing, improved in Detection but deteriorated in Telling, while Crew F, with feedback, improved in Telling but not in Detection.

Crew FD added to the understanding of effects of feedback. For this study, the feedback was quite detailed, not defining problems for the crew but giving facts from which members could induce problems. The leaders of this crew were extremely anxious persons (more so than in any other crew), more inclined to search for an excuse for a failure than to define problems and formulate solutions. In the present instance detailed feedback focused the group's attention upon specific acts. Considerable buck-passing occurred inasmuch as the feedback engendered much hostility. This situation does not negate the observations that feed back can be helpful, particularly if the report is designed for the specific group. This matter is treated at some length in another paper (Tilton & Jensen 1960), Nor does it negate the findings that specific information can lead to considerable improvement in performance.

The general results support the idea that feedback (knowledge of results) and debriefings (discussions) contribute to improvement in crew performance. It is believed that the major way in which these factors contribute is through aiding in the development of a model of the environment, in terms of which goals may be set and procedures worked out. In the absence of such a model, performance is likely to be undirected: the excessively vigorous approach of Crew O is an example of a reaction which was probably as satisfactory as was possible in the circumstances. The desire on the part of subjects for such a model has been indicated by subsequent laboratory studies in which we have noticed that considerable effort is directed towards the question, " What is the nature of the environment ? "

The research reported herein is being performed pursuant to a contact with the United States Air Force, Department of Defense.

Une expérience d'entraînement a été instaurée dans une station de contrôle par radar aux Etats-Unis. Des conditions extérieures et imprévues ont atténué à tel point les conditions expérimentales qu'il a été difficile d'évaluer les effets de façon non ambigüe. Il semble toutefois que la connaissance des résultats et les discussions puissent améliorer les performances dans l'équipe en permettant une meilleure compréhension de l'environnement et des modalités de réponse chez les membres du groupe. Le présent article expose cette expérimentation et interprète les variations de la performance en fonction du genre d'information fourni aux équipes.

Auf einer Radarstation der amerikanischen Luftwaffe wurde ein Trainingsexperiment in Angriff genommen. Äussere und unvorhergesehene Ereignisse beeinflussen die experimentellen Bedingungen so, dass die Wirkungen nicht unzweideutig bewertet werden können. Jedoch unterstützt die Arbeitsveränderung einer Mannschaft die Vorstellung, dass die Kenntnis der Ergebnisse und Diskussionen zur Verbesserung beiträgt, indem sie dabei helfen, die Umgebungsbedingungen und die typischen Reaktionen der Mitglieder einer Mannschaft zu verstehen. Diese Arbeit beschreibt die Studie und bespricht eingehend die Veränderungen in bezug auf die Informationen, über die die Mannschaft verfügte.

References

Alexander, L. T., *et al*, 1959. Problems encountered in developing and maintaining a field system training programme. SP–107, System Development Corporation. Pp 14.

Berkowitz, M. I., 1958, Feedback through the supervisory loop. TM–197, System Development Corporation. Pp. 14.

Chapman, R. L., *et al*, 1959, the Systems research laboratory's air defence experiments. *Management Science* **5**, 250–269

Geddes, E., Kepner, C. H., and Tregoe, B. B., 1959. On the concept of visible and invisible functioning. TM–409, System Development Corporation. Pp. 17.

Goodwin, W. R., 1957, The System Development Corporation and system training. *The American Psychologist*, **12**, 524–528.

Tilton, J. R., and Jensen, B. T., 1960, Facilitating self-evaluation in task-orientated group learning. *Human Factors*, **2**, 92–96.

LETTER TO THE EDITOR

Dear Sir,

Franklin V. Taylor Award

Nominations are now being sought for the recipient of the first Franklin V. Taylor Award for outstanding contributions in the field of engineering psychology. This award was established by Division 21 (Engineering Psychology) of the American Psychological Association in September 1960. It is planned to make the first presentation at the annual meeting of the Division in September 1961. Membership in Division 21 is *not* a prerequisite for the nominee. The deadline for receipt of nominations is 1 July 1961.

I shall be pleased to supply nomination blanks upon request.

Yours etc.,

Leonard C. Mead,
Chairman of the Award Committee.

Tufts University,
Medford 55, Mass.,
U.S.A.

BOOK REVIEWS

Engineering Management. By STRUAN A. ROBERTSON. (Glasgow : Blackie & Sons Ltd., 2nd edition, 1960.) [Pp. xii+460.] 30*s.*

THE purpose of the book is to provide a basis for a course of lectures suitable for degree and Higher National students in preparation for the examination of the Institution of Mechanical Engineers in Industrial Administration. However, even the general reader interested in management, should find much wisdom and inspiration on many pages. Management is sub-divided into three parts: The Art of Management, The Science of Management, and The Practice of Management. The Art of Management emphasises the less tangible attributes for successful management, in particular the initiative required to create an enterprise and to inspire a staff to work with conscious desire for its continuing prosperity. The Science of Management includes chapters dealing with the structure of industry, the broad legal aspects of which a manager should be aware, and the developments of scientific management over the past two hundred years. The Practice of Management forms the bulk of the book. Commencing with the preliminary consideration to be taken before starting a new enterprise, the sections develop logically through policy formulation, company promotion and organization, thence to the factory, the work people, production, the customer and finally to the methods of overall control for checking the well-being of the business. In the section on work people there are short chapters on personnel selection, contracts of employment, T.W.1, technical and management education. Sound treatment is given to the problems of Industrial Health and Safety. Mention is also made of the areas in which Operations Research and Cybernetics can make effective contribution to management. J. CHERRY.

People and Productivity. By A. B. WARING. (London : British Productivity Council, 1959.) [Pp. iv+25.] 25*s.*

THE author, who is Chairman of Joseph Lucas (Industries) Ltd., is at present also Chairman of the British Productivity Council. His association with scientific management is of long standing and he is well known for his liberal approach to the human relations problems of industry. All this is reflected in this typographically excellent booklet. Four important aspects of personnel management are described under the section headings: Factory Competitions, Workpeople's Production Committees, Workpeople's Suggestion Schemes, Support for the Foreman. Each section begins with a short homily which endeavours to set out the thinking behind the procedures described, and concludes with hortative warnings and advice derived, no doubt, from a hard school of experience. But the hard core of the sections, which is called "Administration Methods", is really first rate material. A more practical guide could hardly be found for those approaching these personnel management practices for the first time. The sample procedures have been carefully thought out and are very well presented; they describe in some detail the routine work involved in carrying out the various schemes described. Mr. Waring can very well claim to have attained his declared object: " To demonstrate some practical methods by which the experience of human relations in industry can be applied to the advantage of production and the satisfaction of both the workpeople and the staff concerned with it." SEYMOUR HILLS.

Research Techniques in Human Engineering. By ALPHONSE CHAPANIS. (Baltimore, Md.: The John Hopkins Press, 1959.) [Pp. xi+316.] $6.00.

IN common with a great many of his American colleagues, Dr. Chapanis is always able to demonstrate a high level of competence as an experimental psychologist which can only have developed from considerable capacity, sound training, wide reading and long experience. In short, he is a professional in the best sense of the word. He also has the less common ability to be both lucid and succinct while retaining an individual somewhat colloquial style. We should, therefore, expect great things from this book, particularly since it is the first one of its kind to appear. Other authors have refrained from making the attempt on the grounds that research techniques can only be learned by doing research, so that, on reading this book one grasps the full significance of each sentence in the light of personal experience, or one sees it as a statement which is either incomprehensible or obvious. Dr. Chapanis counters this kind of argument on the grounds that the tactics and strategy of research can be taught. This, of course, is different from teaching the research techniques of a particular subject. However, not only does the author know about the techniques but it emerges very clearly that he has actually used them so that the practical as well as the theoretical points are made. This is the real value of the book. It should convey a great deal to students who have just begun research and to research workers who wish to use a technique which has previously been outside their practice. For example Chapter 2 may do a considerable service for British biologists in persuading them to pay more attention to what we could call Methods Study. These procedures are simple, quick and inexpensive, but still very effective and elegant if properly used. As a book for study by undergraduates most of the material can, perhaps, be justified at the advanced level of evaluating original research and for engineers one can hope that they will be persuaded to be rather more cautious in embarking on experiments in which human behaviour is involved. The chapters alternate between practical/philosophical essays (Introduction, Methods for Study of Accidents, The Experimental Methods, Some Problems of Experimenting with People) well illustrated by experiments from the literature and the more usual test-book material (Methods of Direct Observation, Statistical Methods, The Psychophysical Methods) which will serve as very useful introductions. The last chapter "Articulation Testing Methods" deals with a much more specialized subject, but it does serve to demonstrate by example the detailed care with which a particular topic of research in human engineering has very often been dealt with. Thus the book covers what little we know, to date, about the methods of experimental psychology, although to justify the title a much longer book dealing also with control engineering and physiological techniques would be desirable.

W. T. SINGLETON.

INSTRUCTIONS TO CONTRIBUTORS

1. Articles for publication should be sent to one of the Editors or to any Member of the Editorial Board.

2. Papers must be in English, French or German. Every paper must be accompanied by a brief summary, and contributors are asked if possible to supply summaries in all three languages.

3. Authors should submit a typescript, double-spaced on one side of the paper only. Footnotes should be avoided. Summaries, tables and legends for diagrams should be typed on separate sheets. Authors must ensure that the lay-out of mathematical and other formulae is clear. The typescript must represent the final form in which the author wishes the article to appear. The cost of any alteration in proof other than printers' errors may be charged to the author.

4. Diagrams should be drawn in black ink on white card or tracing paper. They should normally be sufficiently large to allow reduction in printing and the lines should therefore be bold. **All lettering should be up to draughtsmanship standard, suitably drawn in Indian ink to allow for reduction in size.** No charges are made for reproducing tables, diagrams or half-tone illustrations, but diagrams not suitable for reproduction without redrawing may be redrawn at the Author's expense.

5. References in the text should be indicated by author's name followed by the date. They should be listed alphabetically at the end of the paper in the style illustrated by the following examples:

BARTLETT, F. C., 1943, Fatigue following highly skilled work. *Proc. roy. Soc.* B, 131, 247–254.

BEDFORD, T., 1948, *Basic Principles of Ventilation and Heating* (London: H. K. Lewis).

LE GROS CLARK, W. E., 1954, The anatomy of work. In *Symposium on Human Factors in Equipment Design* (Edited by W. F. Floyd and A. T. Welford) (London: H. K. Lewis). Pp. 5–15.

Abbreviations should be as in the *World List of Scientific Periodicals.*

6. **Consideration for publication will gladly be given to papers which have previously had a limited circulation as research reports.** Submission of a paper implies, however, that it has not been published and will not be published elsewhere without the permission of the General Editor and the Publishers. Copyright in material accepted for publication is retained by the Journal, and reproduction in whole or in part is forbidden except under the terms of the *Fair Copying Declaration* of the Royal Society or with the written permission of the Publishers.

7. Authors will receive 25 copies of their contributions without charge. Additional copies may be ordered at the time of returning proofs. Prices for additional copies may be obtained from the Publishers.

8. Authors of papers of ergonomic interest published in other journals and of unpublished reports which are available on request, are invited to submit brief summaries to the Reviews Editor for inclusion in this journal. Summaries must be accompanied by a copy of the full paper.

Books and other material for review should be sent to the Reviews Editor.

ERGONOMICS

Volume 4 April 1961 Number 2

Contents

Page

A Practical Method of Estimating an Individual's Maximal Oxygen Intake. By J. S. Maritz, J. F. Morrison, J. Peter, P. B. Strydom and C. H. Wyndham 97

The Estimation of Body Measurements of British Population in Relation to Seat Design. By D. Barkla 123

Neck Muscle Tension and the Postural Image. By F. P. Jones, F. E. Gray, J. A. Hanson and J. D. Shoop 133

Fatigue and Driving. By A. Crawford 143

Some Experimental Studies of Radar Operating. By D. Wallis and J. A. Samuel 155

Some Characteristics of Instrument Measuring Dials. By Antione Papaloïzos 169

Differential Effects of Knowledge of Results and Discussion. By B. J. Jensen, J. R. Tilton and D. N. Anderson 183

Letter to the Editor 189

Book Reviews 191